



McDonald's (TM)

```
D-      446 SIGTN-DLIQYSFAADRMKNVSY-LYQPLNPALLKLIYLITIEGKKVNDIWTMCGCEMAG 503
          || :| |:| | : :: | || : ||:| |:| :: | : |: ::
Qy      136 SIDSNISLVHYIVASQVWMITRYDLYHNFRPAVLLLMFLSVYKAFVMETFVHLCS-LGS 194
          || :| |:| | : :: | || : ||:| |:| :: | : |: ::
D-      504 -EPL-AIPLLLLGL-GKLEFS 520
```

Qy 195 WARLDARAVVTGLLALKHFG 214

## RESULT 2

ID CYB6\_ORYSA STANDARD; PRT; 215 AA.  
AC P12123;  
DT 01-OCT-1989 (REL. 12, CREATED)  
DT 01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)  
DT 01-NOV-1990 (REL. 16, LAST ANNOTATION UPDATE)  
DE CYTOCHROME B6 (EC 1.10.99.1).  
GN PETB.  
OS ORYZA SATIVA (RICE).  
OG CHLOROPLAST.  
OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; MONOCOTYLEDONEAE;  
OC CYPERALES; GRAMINEAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. NIPPONBARE;  
RA SUGIURA M.;  
RL SUBMITTED (JUL-1989) TO EMBL/GENBANK/DBJ DATA BANKS.  
RN [2]  
RP COMPLETE GENOME.  
RX MEDLINE; 89364698.  
RA HIRATSUKA J., SHIMADA H., WHITTIER R., ISHIBASHI T., SAKAMOTO M.,  
RA MORI M., KONDO C., HONJI Y., SUN C.-R., MENG B.-Y., LI Y.-Q.,  
RA KANNO A., NISHIZAWA Y., HIRAI A., SHINOZAKI K., SUGIURA M.;  
RL MOL. GEN. GENET. 217:185-194(1989).  
RN [3]  
RP SEQUENCE FROM N.A.  
RA COTE J.C., WU N.H., WU R.;  
RL PLANT MOL. BIOL. 11:873-874(1988).  
CC -!- FUNCTION: COMPONENT OF THE CYTOCHROME B6/F COMPLEX WHICH IS PART  
CC OF THE CHLOROPLASTIC RESPIRATORY CHAIN.  
CC -!- CATALYTIC ACTIVITY: PLASTOQUINOL-1 + 2 OXYDIZED PLASTOCYANIN =  
CC PLASTOQUINONE + 2 REDUCED PLASTOCYANIN.  
CC -!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY  
CC BOUND TO THE PROTEIN.  
CC -!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B6-F ARE: CYTOCHROME B6,  
CC 17 KD POLYPEPTIDE (PETD), CYTOCHROME F AND THE RIESKE PROTEIN.  
CC -!- SIMILARITY: WITH THE AMINO END OF MITOCHONDRIAL CYTOCHROME B.  
DR EMBL; X15901; G669082; -.  
DR EMBL; M35995; G343209; -.  
DR PIR; JQ0256; CBR26.  
DR MENDEL; 4313; ORYSA; PETB; 1.  
DR PROSITE; PS00192; CYTOCHROME\_B\_HEME; 1.  
KW ELECTRON TRANSPORT; HEME; CHLOROPLAST; PHOTOSYNTHESIS; TRANSMEMBRANE.  
FT METAL 86 86 IRON 1 (HEME AXIAL LIGAND)  
FT (BY SIMILARITY).  
FT METAL 100 100 IRON 2 (HEME AXIAL LIGAND)  
FT (BY SIMILARITY).  
FT METAL 187 187 IRON 2 (HEME AXIAL LIGAND)  
FT (BY SIMILARITY).  
FT METAL 202 202 IRON 1 (HEME AXIAL LIGAND)  
FT (BY SIMILARITY).  
SQ SEQUENCE 215 AA; 24182 MW; 086A69C5 CRC32;

Query Match 5.8%; Score 113; DB 1; Length 215;  
Best Local Similarity 23.5%; Pred. No. 1.21e-02;  
Matches 20; Conservative 27; Mismatches 32; Indels 6; Gaps 6;

Db 57 YYPRTVTEAFSSVQYIMTEANFGWLI-R-SV-HRWSASMMVLMILHVFVRYLTGGFKKP 113

Qy 132 YIQMSIDSNISLVHYIIVASAQV-WMITRYDLHNFPAVL-LLMFLSVYKAFVMTFVHL 189

Db 114 REL-TWVTGVVLAVLTASFGVTGYS 137

Qy 190 CSLGSWARLDARAVVTGLLALKHFG 214

## RESULT 3

ID CYB6\_MAIZE STANDARD; PRT; 215 AA.  
AC P05642;  
DT 01-NOV-1988 (REL. 09, CREATED)

DT 01-NOV-1988 (REL. 09, LAST SEQUENCE UPDATE)  
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
DE CYTOCHROME B6 (EC 1.10.99.1).  
GN PETB.  
OS ZEA MAYS (MAIZE).  
OG CHLOROPLAST.  
OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; MONOCOTYLEDONEAE;  
OC CYPERALES; GRAMINEAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 88210525.  
RA ROCK C.D., BARKAN A., TAYLOR W.C.;  
RL CURR. GENET. 12:69-77(1987).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 95395841.  
RA MAIER R.M., NECKERMAN K., IGLOI G.L., KOESSEL H.;  
RL J. MOL. BIOL. 251:614-628(1995).  
CC -!- FUNCTION: COMPONENT OF THE CYTOCHROME B6/F COMPLEX WHICH IS PART  
CC OF THE CHLOROPLASTIC RESPIRATORY CHAIN.  
CC -!- CATALYTIC ACTIVITY: PLASTOQUINOL-1 + 2 OXYDIZED PLASTOCYANIN =  
CC PLASTOQUINONE + 2 REDUCED PLASTOCYANIN.  
CC -!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY  
CC BOUND TO THE PROTEIN.  
CC -!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B6-F ARE: CYTOCHROME B6,  
CC 17 KD POLYPEPTIDE (PETD), CYTOCHROME F AND THE RIESKE PROTEIN.  
CC -!- SIMILARITY: WITH THE AMINO END OF MITOCHONDRIAL CYTOCHROME B.  
DR EMBL; X05422; G311718; -.  
DR EMBL; X86563; G902251; ALT\_SEQ.  
DR PIR; S08592; CBZM6R.  
DR MAIZEB; 56334; -.  
DR MENDEL; 4315; ZEAMA; PETB; 1.  
DR PROSITE; PS00192; CYTOCHROME\_B\_HEME; 1.  
KW ELECTRON TRANSPORT; HEME; CHLOROPLAST; PHOTOSYNTHESIS; TRANSMEMBRANE.  
FT METAL 86 86 IRON 1 (HEME AXIAL LIGAND)  
FT (BY SIMILARITY).  
FT METAL 100 100 IRON 2 (HEME AXIAL LIGAND)  
FT (BY SIMILARITY).  
FT METAL 187 187 IRON 2 (HEME AXIAL LIGAND)  
FT (BY SIMILARITY).  
FT METAL 202 202 IRON 1 (HEME AXIAL LIGAND)  
FT (BY SIMILARITY).  
SQ SEQUENCE 215 AA; 24180 MW; CBAB4737 CRC32;

Query Match 5.8%; Score 113; DB 1; Length 215;  
Best Local Similarity 23.5%; Pred. No. 1.21e-02;  
Matches 20; Conservative 27; Mismatches 32; Indels 6; Gaps 6;

Db 57 YYPRTVTEAFSSVQYIMTEANFGWLI-R-SV-HRWSASMMVLMILHVFVRYLTGGFKKP 113

Qy 132 YIQMSIDSNISLVHYIIVASAQV-WMITRYDLHNFPAVL-LLMFLSVYKAFVMTFVHL 189

Db 114 REL-TWVTGVVLAVLTASFGVTGYS 137

Qy 190 CSLGSWARLDARAVVTGLLALKHFG 214

## RESULT 4

ID CYB6\_WHEAT STANDARD; PRT; 215 AA.  
AC P12362;  
DT 01-OCT-1989 (REL. 12, CREATED)  
DT 01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)  
DT 01-DEC-1992 (REL. 24, LAST ANNOTATION UPDATE)  
DE CYTOCHROME B6 (EC 1.10.99.1).  
GN PETB.  
OS TRITICUM AESTIVUM (WHEAT), AND HORDEUM VULGARE (BARLEY).  
OG CHLOROPLAST.  
OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; MONOCOTYLEDONEAE;  
OC CYPERALES; GRAMINEAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC SPECIES=WHEAT; STRAIN=CV. MARDLER;  
RX MEDLINE; 91329710.

```

Db      11 IFMSSLISMILTIY-LNNIFNPSMLLIYLISYSIYMSLMFT-MCSMNSLLILMILIVF 68
      | : : | | : : : : : : : : : : : : : : : : : : : : : : : :
Qy     147 IVASAQWMI-TRYDLYHNFRPAVLLMFLSVYKAFV-METFVHLCSLGSGWARLDARAV- 203

```

```

ID  CYB6_NOSSP          STANDARD;          PRT;    215 AA.
AC  P12122;
DT  01-OCT-1989 (REL. 12, CREATED)
DT  01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)
DT  01-APR-1990 (REL. 14, LAST ANNOTATION UPDATE)
DE  CYTOCHROME B6 (EC 1.10.99.1).
GN  PETB.
OS  NOSTOC SP. (STRAIN PCC 7906).
OC  PROKARYOTA; GRACILICUTES; OXYPHOTOBACTERIA;
OC  CYANOBACTERIA (BLUE-GREEN ALGAE); NOSTOCALES.
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE; 89008280.
RA  KALLAS T., SPILLER S., MALKIN R.;
RL  J. BIOL. CHEM. 263:14334-14342(1988).
CC  -!- FUNCTION: COMPONENT OF THE CYTOCHROME B6/F COMPLEX WHICH IS PART
CC  OF THE CYANOBACTERIAL RESPIRATORY CHAIN.
CC  -!- CATALYTIC ACTIVITY: PLASTOQUINOL-1 + 2 OXYDIZED PLASTOCYANIN =
CC  PLASTOQUINONE + 2 REDUCED PLASTOCYANIN.
CC  -!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC  BOUND TO THE PROTEIN.

```

```

CC      -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B6-F ARE: CYTOCHROME B6,
CC      17 KD POLYPEPTIDE (PETD), CYTOCHROME F AND THE RIESKE PROTEIN.
CC      -1- SIMILARITY: WITH THE AMINO END OF MITOCHONDRIAL CYTOCHROME B.
DR      EMBL; J03967; G145023; -.
DR      PIR; A30807; A30807.
DR      PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
KW      ELECTRON TRANSPORT; HEME; PHOTOSYNTHESIS; TRANSMEMBRANE.
FT      METAL      86      86      IRON 1 (HEME AXIAL LIGAND)
FT                                     (BY SIMILARITY).
FT      METAL      100     100     IRON 2 (HEME AXIAL LIGAND)
FT                                     (BY SIMILARITY).
FT      METAL      187     187     IRON 2 (HEME AXIAL LIGAND)
FT                                     (BY SIMILARITY).
FT      METAL      202     202     IRON 1 (HEME AXIAL LIGAND)
FT                                     (BY SIMILARITY).
SQ      SEQUENCE   215 AA;  24259 MW;  DEAE4993 CRC32;

```

Query Match 5.5%; Score 106; DB 1; Length 215;  
Best Local Similarity 25.0%; Pred. No. 1.07e-01;  
Matches 19; Conservative 18; Mismatches 33; Indels 6; Gaps 6;

```

Db      57 YYKPTVAEAFSSVEYIMNEVFGWLI-R-SI-HRWSASMMVLMMLIHVFRVYLGGFKPK 113
      | :: :| ||: |::| :| :::: |:|:| |: :: |
Qy     132 YIQMSIDSNISLVHYIVASQV-WMITRDLYHNFRPAVL-LLMFLSVYKAFVMTTFVHL 189

Db     114 REL-TWVSGVILAVIT 128
      * | :| ||:|
Qy     190 CSLGSWARLDARAVVT 205

```

RESULT 10  
ID CYB6\_SPIOL STANDARD; PRT: 215 AA.  
AC P00165;  
DT 21-JUL-1986 (REL. 01, CREATED)  
DT 01-NOV-1988 (REL. 09, LAST SEQUENCE UPDATE)  
DT 01-APR-1990 (REL. 14, LAST ANNOTATION UPDATE)  
DE CYTOCHROME B6 (EC 1.10.99.1).  
GN PETB.  
OS SPINACIA OLERACEA (SPINACH).  
OC CHLOROPLAST.  
OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;  
OC CARYOPHYLLALES; CHENOPODIACEAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 88194670.  
RA WESTHOFF P., FARCHAU J.W., HERRMANN R.G.;  
RL CURR. GENET. 11:165-169(1986).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA HEINEMEYER W., ALT J., HERRMANN R.G.;  
RL CURR. GENET. 8:543-549(1984).  
RN [3]  
RP SEQUENCE.  
RX MEDLINE; 84144766.  
RA WIDGER W.R., CRAMER W.A., HERRMANN R.G., TREBST A.;  
RL PROC. NATL. ACAD. SCI. U.S.A. 81:674-678(1984).  
CC -1- COFACTOR: COMPONENT OF THE CYTOCHROME B6/F COMPLEX WHICH IS PART  
CC OF THE CHLOROPLASTIC RESPIRATORY CHAIN.  
CC -1- CATALYTIC ACTIVITY: PLASTOQUINOL-1 + 2 OXYDIZED PLASTOCYANIN =  
CC PLASTOQUINONE + 2 REDUCED PLASTOCYANIN.  
CC -1- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY  
CC BOUND TO THE PROTEIN.  
CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B6-F ARE: CYTOCHROME B6,  
CC 17 KD POLYPEPTIDE (PETD), CYTOCHROME F AND THE RIESKE PROTEIN.  
CC -1- SIMILARITY: WITH THE AMINO END OF MITOCHONDRIAL CYTOCHROME B.  
DR EMBL; X07106; G12285; -.  
DR PIR; A00161; CBSP6.  
DR PIR; S00429; S00429.  
DR MENDEL; 2553; SPIOL; PETB; 1.  
DR PROSITE; PS00192; CYTOCHROME\_B\_HEME; 1.  
KW ELECTRON TRANSPORT; HEME; CHLOROPLAST; PHOTOSYNTHESIS; TRANSMEMBRANE.  
FT METAL 86 86 IRON 1 (HEME AXIAL LIGAND).  
FT METAL 100 100 IRON 2 (HEME AXIAL LIGAND).

```

FT METAL      187      187      IRON 2 (HEME AXIAL LIGAND).
FT METAL...   202      202      IRON-1 (HEME AXIAL LIGAND).
FT CONFLICT   108      109      MISSING (IN REF. 3).
FT CONFLICT   188      188      MISSING (IN REF. 3).
SQ SEQUENCE   215 AA;  24166 MW;  35E50AD CRC32;

```

Query Match 5.5%; Score 106; DB 1; Length 215;  
Best Local Similarity 22.7%; Pred. No. 1.07e-01;  
Matches 15; Conservative 21; Mismatches 24; Indels 6; Gaps 6;

Db 57 YYRPTVTDAFASVQYIMTEVNFGLWI-R-SV-HRWSASMMVLMILHVFVRYLTGGFKKP 113  
| : :: :: |:|:: |:| : | :::: |:| :| :: :: |  
Qy 132 YIQMSIDSNSISLVHYIVASAQV-WMITRDLYHNFRPAVL-LLMFLSVYKAFVMTFVHL 189

```

Db      114 REL-TW 118
        | :|
Qy      190 CSLGSW 195

```

```

RESULT 11
ID  CYB6_MARPO      STANDARD;      PRT;    215 AA.
AC  P06248;
DT  01-JAN-1988 (REL. 06, CREATED)
DT  01-JAN-1988 (REL. 06, LAST SEQUENCE UPDATE)
DT  01-MAR-1992 (REL. 21, LAST ANNOTATION UPDATE)
DE  CYTOCHROME B6 (EC 1.10.99.1).
GN  PETB.
OS  MARCHANTIA POLYMORPHA (LIVERWORT).
OG  CHLOROPLAST.
OC  EUKARYOTA; PLANTA; EMBRYOPHYTA; BRYOPHYTA; HEPATICOPSIDA.
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE; 89068687.
RA  FUKUZAWA H., KOHCHI T., SANO T., SHIRAI H., UMESONO K., INOKUCHI H.,
RA  OZEKI H., OHYAMA K.;
RL  J. MOL. BIOL. 203:333-351(1988).
RN  [2]
RP  COMPLETE GENOME.
RA  OHYAMA K., FUKUZAWA H., KOHCHI T., SHIRAI H., SANO T., SANO S.,
RA  UMESONO K., SHIKI Y., TAKEUCHI M., CHANG Z., AOTA S., INOKUCHI H.,
RA  OZEKI H.;
RL  NATURE 322:572-574(1986).
RN  [3]
RP  SEQUENCE FROM N.A.
RA  FUKUZAWA H., YOSHIDA T., KOHCHI T., OKUMURA T., SAWANO Y., OHYAMA K.;
RA  FEBS LETT. 220:61-66(1987).
CC  -1- FUNCTION: COMPONENT OF THE CYTOCHROME B6/F COMPLEX WHICH IS PART
CC  OF THE CHLOROPLASTIC RESPIRATORY CHAIN.
CC  -1- CATALYTIC ACTIVITY: PLASTOQUINOL-1 + 2 OXYDIZED PLASTOCYANIN =
CC  PLASTOQUINONE + 2 REDUCED PLASTOCYANIN.
CC  -1- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC  BOUND TO THE PROTEIN.
CC  -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B6-F ARE: CYTOCHROME B6,
CC  17 KD POLYPEPTIDE (PETD), CYTOCHROME F AND THE RIESKE PROTEIN.
CC  -1- SIMILARITY: WITH THE AMINO END OF MITOCHONDRIAL CYTOCHROME B.
DR  EMBL; X04465; E122080; -.
DR  PIR; A00163; CBLV6.
DR  PIR; S01552; S01552.
DR  PIR; S02432; S02432.
DR  MENDEL; 2551; MARPO;PETB;1.
DR  PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
KW  ELECTRON TRANSPORT; HEME; CHLOROPLAST; PHOTOSYNTHESIS; TRANSMEMBRANE.
FT  METAL      86      86      IRON 1 (HEME AXIAL LIGAND)
FT                                     (BY SIMILARITY).
FT  METAL      100     100     IRON 2 (HEME AXIAL LIGAND)
FT                                     (BY SIMILARITY).
FT  METAL      187     187     IRON 2 (HEME AXIAL LIGAND)
FT                                     (BY SIMILARITY).
FT  METAL      202     202     IRON 1 (HEME AXIAL LIGAND)
FT                                     (BY SIMILARITY).
SQ  SEQUENCE    215 AA;  24307 MW;  A8EDE412 CRC32;

```

Query Match 5.4%; Score 105; DB 1; Length 215;

Best Local Similarity 22.7%; Pred. No. 1.45e-01;  
Matches 15; Conservative 21; Mismatches 24; Indels 6; Gaps 6;

Db 57 YYRPTVTEAFSSVQYIMTEVNFGLI-R-SV-HRWSASMMVLMILHIFRVYLTGGFKKP 113  
| : : : : | : : : : | : | : | : : : | : : : : |  
Qy 132 YIQMSIDSNISLVHYIVASAQV-WMITRYDLYHNFRPAVL-LLMFLSVYKAFVMTFVHL 189

Db 114 REL-TW 118

Qy 190 CSLGSW 195

## RESULT 12

ID CYB6\_ODOSI STANDARD; PRT; 215 AA.  
AC P49488;  
DT 01-FEB-1996 (REL. 33, CREATED)  
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)  
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
DE CYTOCHROME B6 (EC 1.10.99.1).  
GN PETB.  
OS ODOMELLA SINENSIS.  
OC CHLOROPLAST.  
OC EUKARYOTA; PLANTA; PHYCOPHYTA; BACILLARIOPHYTA (DIATOMS).  
RN [1]  
RP SEQUENCE FROM N.A.  
RA KOWALLIK K.V., STOEBE B., SCHAFFRAN I., KROTH-PANCIC P., FREIER U.;  
RL PLANT MOL. BIOL. REP. 13:336-342(1995).  
CC -!- FUNCTION: COMPONENT OF THE CYTOCHROME B6/F COMPLEX WHICH IS PART  
CC OF THE CHLOROPLASTIC RESPIRATORY CHAIN.  
CC -!- CATALYTIC ACTIVITY: PLASTOQUINOL-1 + 2 OXYDIZED PLASTOCYANIN =  
CC PLASTOQUINONE + 2 REDUCED PLASTOCYANIN.  
CC -!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY  
CC BOUND TO THE PROTEIN.  
CC -!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B6-F ARE: CYTOCHROME B6,  
CC 17 KD POLYPEPTIDE (PETD), CYTOCHROME F AND THE RIESKE PROTEIN.  
CC -!- SIMILARITY: WITH THE AMINO END OF MITOCHONDRIAL CYTOCHROME B.  
DR EMBL; Z67753; E211859; -.  
DR MENDEL; 2552; ODOSI; PETB; 1.  
DR PROSITE; PS00192; CYTOCHROME\_B\_HEME; FALSE\_NEG.  
KW ELECTRON TRANSPORT; HEME; CHLOROPLAST; PHOTOSYNTHESIS; TRANSMEMBRANE.  
FT METAL 86 86 IRON 1 (HEME AXIAL LIGAND)  
FT (BY SIMILARITY).  
FT METAL 100 100 IRON 2 (HEME AXIAL LIGAND)  
FT (BY SIMILARITY).  
FT METAL 187 187 IRON 2 (HEME AXIAL LIGAND)  
FT (BY SIMILARITY).  
FT METAL 202 202 IRON 1 (HEME AXIAL LIGAND)  
FT (BY SIMILARITY).  
SQ SEQUENCE 215 AA; 23923 MW; 79C43185 CRC32;

Query Match 5.4%; Score 104; DB 1; Length 215;  
Best Local Similarity 26.3%; Pred. No. 1.97e-01;  
Matches 20; Conservative 19; Mismatches 31; Indels 6; Gaps 6;

Db 57 YYRPSVVDFAFSSVEYIMTSVNFGLI-R-SI-HRWSASMMVLMVHFRVYLTGGFKKP 113  
| : : : : | : : : : | : | : | : : : | : : : : |  
Qy 132 YIQMSIDSNISLVHYIVASAQV-WMITRYDLYHNFRPAVL-LLMFLSVYKAFVMTFVHL 189

Db 114 REL-TWTVGVILAVVT 128

Qy 190 CSLGSWRLDARAVVT 205

## RESULT 13

ID CYB6\_PORPU STANDARD; PRT; 215 AA.  
AC P51341;  
DT 01-OCT-1996 (REL. 34, CREATED)  
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
DE CYTOCHROME B6 (EC 1.10.99.1).  
GN PETB.  
OS PORPHYRA PURPUREA.  
OC CHLOROPLAST.

OC EUKARYOTA; PLANTA; PHYCOPHYTA; RHODOPHYTA (RED ALGAE).  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-AVONPORT;  
RA REITH M.E., MUNHOLLAND J.;  
RL PLANT MOL. BIOL. REP. 13:333-335(1995).  
CC -!- FUNCTION: COMPONENT OF THE CYTOCHROME B6/F COMPLEX WHICH IS PART  
CC OF THE CHLOROPLASTIC RESPIRATORY CHAIN.  
CC -!- CATALYTIC ACTIVITY: PLASTOQUINOL-1 + 2 OXYDIZED PLASTOCYANIN =  
CC PLASTOQUINONE + 2 REDUCED PLASTOCYANIN.  
CC -!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY  
CC BOUND TO THE PROTEIN.  
CC -!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B6-F ARE: CYTOCHROME B6,  
CC 17 KD POLYPEPTIDE (PETD), CYTOCHROME F AND THE RIESKE PROTEIN.  
CC -!- SIMILARITY: WITH THE AMINO END OF MITOCHONDRIAL CYTOCHROME B.  
DR EMBL; U38804; G1276807; -.  
DR PROSITE; PS00192; CYTOCHROME\_B\_HEME; 1.  
KW ELECTRON TRANSPORT; HEME; CHLOROPLAST; PHOTOSYNTHESIS; TRANSMEMBRANE.  
FT METAL 86 86 IRON 1 (HEME AXIAL LIGAND)  
FT (BY SIMILARITY).  
FT METAL 100 100 IRON 2 (HEME AXIAL LIGAND)  
FT (BY SIMILARITY).  
FT METAL 187 187 IRON 2 (HEME AXIAL LIGAND)  
FT (BY SIMILARITY).  
FT METAL 202 202 IRON 1 (HEME AXIAL LIGAND)  
FT (BY SIMILARITY).  
SQ SEQUENCE 215 AA; 24224 MW; 9F90E552 CRC32;

Query Match 5.4%; Score 104; DB 1; Length 215;

Best Local Similarity 22.7%; Pred. No. 1.97e-01;  
Matches 15; Conservative 20; Mismatches 25; Indels 6; Gaps 6;

Db 57 YYRPTVAEAFSSVEYIMTDVNFGLI-R-SI-HRWSASMMVLMILHIFRVYLTGGFKKP 113  
| : : : : | : : : : | : | : | : : : | : : : : |  
Qy 132 YIQMSIDSNISLVHYIVASAQV-WMITRYDLYHNFRPAVL-LLMFLSVYKAFVMTFVHL 189

Db 114 REL-TW 118

Qy 190 CSLGSW 195

## RESULT 14

ID CYB6\_SYN2 STANDARD; PRT; 222 AA.  
AC P28056;  
DT 01-AUG-1992 (REL. 23, CREATED)  
DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)  
DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)  
DE CYTOCHROME B6 (EC 1.10.99.1).  
GN PETB.  
OS SYNECHOCOCCUS SP. (STRAIN PCC 7002) (AGMENELLUM QUADRUPLICATUM).  
OC PROKARYOTA; GRACILICUTES; OXYPHOTOBACTERIA;  
OC CYANOBACTERIA (BLUE-GREEN ALGAE); CHROCOCCALES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 93043038.  
RA BRAND S.N., TAN X., WIDGER W.R.;  
RL PLANT MOL. BIOL. 20:481-491(1992).  
CC -!- FUNCTION: COMPONENT OF THE CYTOCHROME B6/F COMPLEX WHICH IS PART  
CC OF THE CYANOBACTERIAL RESPIRATORY CHAIN.  
CC -!- CATALYTIC ACTIVITY: PLASTOQUINOL-1 + 2 OXYDIZED PLASTOCYANIN =  
CC PLASTOQUINONE + 2 REDUCED PLASTOCYANIN.  
CC -!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY  
CC BOUND TO THE PROTEIN.  
CC -!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B6-F ARE: CYTOCHROME B6,  
CC 17 KD POLYPEPTIDE (PETD), CYTOCHROME F AND THE RIESKE PROTEIN.  
CC -!- SIMILARITY: WITH THE AMINO END OF MITOCHONDRIAL CYTOCHROME B.  
DR EMBL; X63049; G38963; -.  
DR PIR; S18123; S18123.  
DR PIR; S26193; S26193.  
DR PROSITE; PS00192; CYTOCHROME\_B\_HEME; 1.  
KW ELECTRON TRANSPORT; HEME; PHOTOSYNTHESIS; TRANSMEMBRANE.  
FT METAL 93 93 IRON 1 (HEME AXIAL LIGAND)  
FT (BY SIMILARITY).







RESULT 2  
ENTRY CBZM6R #type complete  
TITLE plastoquinol--plastocyanin reductase (EC 1.10.99.1)  
cytochrome b6 - maize chloroplast  
ORGANISM #formal\_name chloroplast Zea mays #common\_name maize  
DATE 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change  
05-Sep-1997  
ACCESSIONS S08592  
REFERENCE S07171  
#authors Rock, C.D.; Barkan, A.; Taylor, W.C.  
#journal Curr. Genet. (1987) 12:69-77  
#title The maize plastid psbB-psbF-petB-petD gene cluster: spliced  
and unspliced petB and petD RNAs encode alternative  
products.  
#cross-references MUID:88210525  
#accession S08592  
##molecule\_type DNA  
##residues 1-215 ##label ROC  
##cross-references EMBL:X05422; NID:g12434; PID:g311718

GENETICS  
#gene petB  
#genome chloroplast  
#introns 2/3  
CLASSIFICATION #superfamily cytochrome b6; cytochrome b6 homology  
KEYWORDS chloroplast; chromoprotein; electron transfer; heme; iron;  
oxidoreductase; photosynthesis; thylakoid  
FEATURE  
15-215 #domain cytochrome b6 homology #label CB6\  
86,202 #binding\_site heme iron (His) (axial ligands) #status  
predicted\  
100,187 #binding\_site heme iron (His) (axial ligands) #status  
predicted  
SUMMARY #length 215 #molecular-weight 24180 #checksum 9478  
Query Match 5.8%; Score 113; DB 1; Length 215;  
Best Local Similarity 23.5%; Pred. No. 8.32e-02;  
Matches 20; Conservative 27; Mismatches 32; Indels 6; Gaps 6;

Db 57 YRPTVTEAFSSVQYIMTEANFGWLI-R-SV-HRWSASMMVLMILHVFVRVLTGGFKPK 113  
| : : : | : | : : | : | : : | : : : | : | : : | : : : |  
Qy 132 YIQMSIDSNISLVHIIVASAQV-WMITRYDLYHNFRAVL-LLMFLSVYKAFVMTFVHL 189  
Db 114 REL-TWVTGVVLAULTASFGVTGYS 137  
| : | | : | : : : :  
Qy 190 CSLGSWARLDARAVVTGLLALKHFG 214

RESULT 3  
ENTRY S14961 #type complete  
TITLE plastoquinol--plastocyanin reductase (EC 1.10.99.1)  
cytochrome b6 - wheat chloroplast  
ALTERNATE\_NAMES cytochrome b563  
ORGANISM #formal\_name chloroplast; Triticum aestivum #common\_name  
common wheat  
DATE 08-Jun-1994 #sequence\_revision 03-Aug-1995 #text\_change  
24-Apr-1998  
ACCESSIONS S14961  
REFERENCE S14960  
#authors Hird, S.M.; Wilson, R.J.; Dyer, T.A.; Gray, J.C.  
#journal Plant Mol. Biol. (1991) 16:745-747  
#title Nucleotide sequence of the wheat chloroplast petB and petD  
genes encoding apocytochrome b-563 and subunit IV of the  
cytochrome b6 complex.  
#cross-references MUID:91329710  
#accession S14961  
##status preliminary  
##molecule\_type DNA  
##residues 1-215 ##label HIR  
##cross-references EMBL:X54751; NID:g12361; PID:g12363  
GENETICS  
#gene petB

#genome chloroplast  
#introns 2/3  
CLASSIFICATION #superfamily cytochrome b6; cytochrome b6 homology  
KEYWORDS chloroplast; oxidoreductase; thylakoid  
FEATURE  
15-215 #domain cytochrome b6 homology #label CB6  
SUMMARY #length 215 #molecular-weight 24166 #checksum 9436  
Query Match 5.8%; Score 113; DB 2; Length 215;  
Best Local Similarity 23.5%; Pred. No. 8.32e-02;  
Matches 20; Conservative 27; Mismatches 32; Indels 6; Gaps 6;  
Db 57 YRPTVTEAFSSVQYIMTEANFGWLI-R-SV-HRWSASMMVLMILHVFVRVLTGGFKPK 113  
| : : : | : | : : | : | : : | : : : | : | : : | : : : |  
Qy 132 YIQMSIDSNISLVHIIVASAQV-WMITRYDLYHNFRAVL-LLMFLSVYKAFVMTFVHL 189  
Db 114 REL-TWVTGVVLAULTASFGVTGYS 137  
| : | | : | : : : :  
Qy 190 CSLGSWARLDARAVVTGLLALKHFG 214

RESULT 4  
ENTRY CBRZ6 #type complete  
TITLE plastoquinol--plastocyanin reductase (EC 1.10.99.1)  
cytochrome b6 - rice chloroplast  
ORGANISM #formal\_name chloroplast Oryza sativa #common\_name rice  
DATE 31-Mar-1990 #sequence\_revision 31-Mar-1990 #text\_change  
18-Jul-1997  
ACCESSIONS JQ0256; S05136  
REFERENCE JQ0200  
#authors Shimada, H.; Whittier, R.F.; Hiratsuka, J.; Maeda, Y.; Hirai,  
A.; Sugiura, M.  
#submission submitted to JIPID, December 1989  
#accession JQ0256  
##molecule\_type DNA  
##residues 1-215 ##label SHI  
##experimental\_source cv. Nihonbare  
REFERENCE S05080  
#authors Hiratsuka, J.; Shimada, H.; Whittier, R.; Ishibashi, T.;  
Sakamoto, M.; Mori, M.; Kondo, C.; Honji, Y.; Sun, C.R.;  
Meng, B.Y.; Li, Y.Q.; Kanno, A.; Nishizawa, Y.; Hirai, A.;  
Shinozaki, K.; Sugiura, M.  
#journal Mol. Gen. Genet. (1989) 217:185-194  
#title The complete sequence of the rice (Oryza sativa) chloroplast  
genome: intermolecular recombination between distinct trna  
genes accounts for a major plastid DNA inversion during the  
evolution of the cereals.  
#cross-references MUID:89364698  
#accession S05136  
##status nucleic acid sequence not shown; translation not shown  
##molecule\_type DNA  
##residues 1-215 ##label HIR  
##cross-references EMBL:X15901; NID:g11957; PID:g669082  
##experimental\_source cv. Nihonbare  
COMMENT This cytochrome is one of the components of a specific  
stoichiometric cytochrome b6-f complex that contains two  
molecules of cytochrome b6, one cytochrome f and one nonheme  
iron-sulfur center.

GENETICS  
#gene petB  
#map\_position CP1232-71237,72049-72690  
#genome chloroplast  
#introns 2/3  
CLASSIFICATION #superfamily cytochrome b6; cytochrome b6 homology  
KEYWORDS chloroplast; chromoprotein; electron transfer; heme; iron;  
oxidoreductase; photosynthesis; thylakoid  
FEATURE  
15-215 #domain cytochrome b6 homology #label CB6\  
86,202 #binding\_site heme iron (His) (axial ligands) #status  
predicted\  
100,187 #binding\_site heme iron (His) (axial ligands) #status  
predicted  
SUMMARY #length 215 #molecular-weight 24182 #checksum 9304

Db 76 YYRPTVTEAFSSVOYIMTEANFGWLI-R-SV-HRWSASMMVLMMILHVERVYLTGGFKKP 132

Oy 132 YIQMSIDSNISLVHYIVASAQV-WMITRYDLYHNFRPAVL-LLMFLSVYKAFVMTFVHL 189  
Db 133 REL-TWVTGVVLAVLTASFGVTGYS 156  
Oy 190 CSLGSWARLDARAVVTGLLALKHFG 214

RESULT 8  
ENTRY CBNT6 #type complete  
TITLE plastoquinol--plastocyanin reductase (EC 1.10.99.1)  
cytochrome b6 - common tobacco chloroplast  
ALTERNATE\_NAMES cytochrome b563  
ORGANISM #formal\_name chloroplast Nicotiana tabacum #common\_name  
common tobacco  
DATE 30-Jun-1987 #sequence\_revision 30-Jun-1987 #text\_change  
04-Oct-1996  
ACCESSIONS A00162  
REFERENCE A00149  
#authors Sugiura, M.  
#submission submitted to the EMBL Data Library, August 1986  
#accession A00162  
##molecule\_type DNA  
##residues 1-215 ##label SUG  
##experimental\_source cv. Bright Yellow 4  
REFERENCE A38013  
#authors Shinozaki, K.; Ohme, M.; Tanaka, M.; Wakasugi, T.; Hayashida,  
N.; Matsubayashi, T.; Zaita, N.; Chunwongse, J.; Obokata,  
J.; Yamaguchi-Shinozaki, K.; Ohto, C.; Torazawa, K.; Meng,  
B.Y.; Sugita, M.; Deno, H.; Kamogashira, T.; Yamada, K.;  
Kusuda, J.; Takaiwa, F.; Kato, A.; Tohdoh, N.; Shimada, H.;  
Sugiura, M.  
#journal EMBO J. (1986) 5:2043-2049  
#title The complete nucleotide sequence of the tobacco chloroplast  
genome: its gene organization and expression.  
#contents annotation; gene organization, sites, features  
COMMENT This cytochrome is one of the components of a specific  
stoichiometric cytochrome b6-f complex that contains two  
molecules of cytochrome b6, one cytochrome f, and one nonheme  
iron-sulfur center.  
GENETICS  
#gene petB  
#genome chloroplast  
#introns 2/3  
CLASSIFICATION #superfamily cytochrome b6; cytochrome b6 homology  
KEYWORDS chloroplast; chromoprotein; electron transfer; heme; iron;  
oxidoreductase; photosynthesis; thylakoid  
FEATURE  
15-215 #domain cytochrome b6 homology #label CB6  
86,202 #binding\_site heme iron (His) (axial ligands) #status  
predicted  
100,187 #binding\_site heme iron (His) (axial ligands) #status,  
predicted  
SUMMARY #length 215 #molecular-weight 24136 #checksum 9633  
Query Match 5.7%; Score 111; DB 1; Length 215;  
Best Local Similarity 22.4%; Pred. No. 1.44e-01;  
Matches 19; Conservative 28; Mismatches 32; Indels 6; Gaps 6;

Db 57 YYRPTVTAEAFASVQYIMTEANFGWLI-R-SV-HRWSASMMVLMMLHVFVRVLTGGFKPK 113  
Oy 132 YIQMSIDSNISLVHYIVASAQV-WMITRYDLYHNFRPAVL-LLMFLSVYKAFVMTFVHL 189  
Db 114 REL-TWVTGVVLAVLTASFGVTGYS 137  
Oy 190 CSLGSWARLDARAVVTGLLALKHFG 214

RESULT 9  
ENTRY S52970 #type complete  
TITLE NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 6 -  
honeybee mitochondrion (SGC4)  
ORGANISM #formal\_name mitochondrion Apis mellifera #common\_name

honeybee  
DATE 15-Jul-1995 #sequence\_revision 10-Nov-1995 #text\_change  
21-Aug-1998  
ACCESSIONS S52970  
REFERENCE S52960  
#authors Crozier, R.H.; Crozier, Y.C.  
#journal Genetics (1993) 133:97-117  
#title The mitochondrial genome of the honeybee Apis mellifera:  
complete sequence and genome organization.  
#accession S52970  
##molecule\_type DNA  
##residues 1-167 ##label CRO  
##cross-references EMBL:L06178; NID:g336279; PID:g829009  
##experimental\_source ligustica  
GENETICS  
#genome mitochondrion  
#genetic\_code SGC4  
#start\_codon ATT  
CLASSIFICATION #superfamily NADH dehydrogenase (ubiquinone) chain 6  
KEYWORDS membrane-associated complex; mitochondrion; NAD;  
oxidoreductase  
SUMMARY #length 167 #molecular-weight 19971 #checksum 6545  
Query Match 5.5%; Score 106; DB 2; Length 167;  
Best Local Similarity 29.9%; Pred. No. 5.58e-01;  
Matches 20; Conservative 17; Mismatches 25; Indels 5; Gaps 5;  
Db 11 IFMSSLISMLTIY-LNNIFNPSMLLIYLSYSIYMSLMFT-MCSMNSLLILMILIVE 68  
Oy 147 IVASQVQWMI-TRYDLYHNFRPAVLLMFLSVYKAFV-METFVHLCSLGSWARLDARAV- 203  
Db 69 LSGMLIM 75  
Oy 204 VTGLLAL 210

RESULT 10  
ENTRY S21253 #type complete  
TITLE plastoquinol--plastocyanin reductase (EC 1.10.99.1)  
cytochrome b6 - Chlamydomonas reinhardtii chloroplast  
ORGANISM #formal\_name chloroplast Chlamydomonas reinhardtii  
DATE 03-Feb-1994 #sequence\_revision 02-Aug-1994 #text\_change  
05-Sep-1997  
ACCESSIONS S21253; S16917  
REFERENCE S20938  
#authors Huang, C.; Liu, X.Q.  
#journal Plant Mol. Biol. (1992) 18:985-988  
#title Nucleotide sequence of the frxC, petB and trnL genes in the  
chloroplast genome of Chlamydomonas reinhardtii.  
#cross-references MUID:92256821  
#accession S21253  
##status translation not shown  
##molecule\_type DNA  
##residues 1-215 ##label HUA  
##cross-references EMBL:X62905; NID:g12497; PID:g12499  
REFERENCE S16916  
#authors Bueschlen, S.; Choquet, Y.; Kuras, R.; Wollman, F.A.  
#journal FEBS Lett. (1991) 284:257-262  
#title Nucleotide sequences of the continuous and separated petA,  
petB and petD chloroplast genes in Chlamydomonas  
reinhardtii.  
#cross-references MUID:91285146  
#accession S16917  
##molecule\_type DNA  
##residues 1-215 ##label BUE  
##cross-references EMBL:X72918; NID:g603530; PID:g288909  
GENETICS  
#gene petB  
#genome chloroplast  
CLASSIFICATION #superfamily cytochrome b6; cytochrome b6 homology  
KEYWORDS chloroplast; chromoprotein; electron transfer; heme; iron;  
oxidoreductase; photosynthesis; thylakoid  
FEATURE

```

15-215      #domain cytochrome b6 homology #label CB6\
86,187      #binding_site heme iron, low potential (His) (axial...
            ligands) #status predicted\
100,202     #binding_site heme iron, high potential (His) (axial
            ligands) #status predicted
SUMMARY     #length 215 #molecular-weight 24164 #checksum 8607

```

Query Match 5.5%; Score 107; DB 1; Length 215;  
Best Local Similarity 22.7%; Pred. No. 4.27e-01;  
Matches 15; Conservative 21; Mismatches 24; Indels 6; Gaps 6;

Db 57 YYPRTVAEAFASVQYIMTDVNFGLWI-R-SI-HRWSASMMVLMMVLHVRVRYLTGGFKRP 113  
| : : : : | : | : : | : | : : | : : : : | :  
Qy 132 YIQMSIDSNISLVHYIVASAQV-WMITRYDLYHNFRPAVL-LLMFLSVYKAFVMTFVHL 189

Db 114 REL-TW 118  
| : |

Qy 190 CSLGSW 195

```

RESULT 11
ENTRY . A30807 #type complete
TITLE plastoquinol--plastocyanin reductase (EC 1.10.99.1)
        cytochrome b6 - Nostoc sp.
ORGANISM #formal_name Nostoc sp.
DATE 01-Dec-1989 #sequence_revision 02-Aug-1994 #text_change
        04-Oct-1996
ACCESSIONS A30807
REFERENCE A94683
#authors Kallas, T.; Spiller, S.; Malkin, R.
#journal J. Biol. Chem. (1988) 263:14334-14342
#title Characterization of two operons encoding the cytochrome b-6-f
        complex of the cyanobacterium Nostoc PCC 7906. Highly
        conserved sequences but different gene organization than in
        chloroplasts.
#cross-references MUID:89008280
#accession A30807
##molecule_type DNA
##residues 1-215 ##label KAL
##cross-references GB:J02961
##experimental_source PCC 7906

```

CLASSIFICATION #superfamily cytochrome b6; cytochrome b6 homology  
KEYWORDS chromoprotein; electron transfer; heme; iron; oxidoreductase;  
                  photosynthesis; thylakoid

```
FEATURE
15-215      #domain cytochrome b6 homology #label CB6\
86,187      #binding_site heme iron, low potential (His) (axial
            ligands) #status predicted\
100,202     #binding_site heme iron, high potential (His) (axial
            ligands) #status predicted
```

SUMMARY #length 215 #molecular-weight 24259 #checksum 8311

Query Match 5.5%; Score 106; DB 1; Length 215;  
Best Local Similarity 25.0%; Pred. No. 5.58e-01;  
Matches 19; Conservative 18; Mismatches 33; Indels 6; Gaps 6;

```

Db . 57 YYKPTVAEAFSSVEYIMNEVNFGLWLR-SI-HRWSASMMVLMMILHVFVRYLTGGFKKP 113
      |  ::  :| |||  :| :| :|  :::: |:| :| :| :|
Qy 132 YIQMSIDSNISLVHYIVASAQV-WMITRYDLYHNFRPAVL-LLMFLSVYKAFVWETFVHL 189

```

Db 114 REL-TWVSGVILAVIT 128  
| : | | |

Qy 190 CSLGSWARLDARAVVT 205

```

RESULT 12
ENTRY      CBSP6          #type complete
TITLE      plastoquinol--plastocyanin reductase (EC 1.10.99.1)
            cytochrome b6 - spinach chloroplast
ORGANISM    #formal_name chloroplast Spinacia oleracea #common_name
            spinach
DATE        03-Aug-1984 #sequence_revision 02-Aug-1994 #text_change
            17-Apr-1998

```

```

ACCESSIONS      S03021; A00161; S00429
REFERENCE..... S03021
#authors        Westhoff, P.; Herrmann, R.G.
#journal        Eur. J. Biochem. (1988) 171:551-564
#title          Complex RNA maturation in chloroplasts.
#cross-references MUID:88151952
#accession      S03021
##molecule_type DNA
##residues      1-22 ##label WES
##note          this sequence is a revision to the sequence from
                reference S00429
##note          not compared with conceptual translation
REFERENCE        A00161
#authors        Widger, W.R.; Cramer, W.A.; Herrmann, R.G.; Trebst, A.
#citation        unpublished results, 1984, cited by Widger, W.R., Cramer,
                W.A., Herrmann, R.G., Trebst, A., Proc. Natl. Acad. Sci.
                U.S.A. 81, 674-678, 1984
#cross-references MUID:84144766
#accession..... A00161

```

```
##status      nucleic acid sequence not shown
##molecule_type DNA
##residues    'MIGSKNVSRFRRLRMI', 21-215 ##label WID
REFERENCE     S00429
#authors      Heinemeyer, W.; Alt, J.; Herrmann, R.G.
#journal      Curr. Genet. (1984) 8:543-549
#title        Nucleotide sequence of the clustered genes for apocytochrome
              b6 and subunit 4 of the cytochrome b/f complex in the
              spinach plastid chromosome.
```

```
#accession      S00429
#molecule_type DNA
#residues       'MIGSKNVSRFRRLRMI',21-215 ##label HEI
#cross-references EMBL:X07106
#note           this sequence has been revised in reference S03021
COMMENT         This cytochrome is the chloroplast counterpart of the mitochondrial
                cytochrome b.
```

GENETICS  
#gene petB  
#genome chloroplast  
#introns 2/3  
CLASSIFICATION #superfamily cytochrome b6; cytochrome b6 homology  
KEYWORDS chloroplast; chromoprotein; electron transfer; heme; iron;  
oxidoreductase; photosynthesis; thylakoid

```

FEATURE
15-215      #domain cytochrome b6 homology #label CB6\
86,187      #binding_site heme iron, low potential (His) (axial
            ligands) #status predicted\
100,202     #binding_site heme iron, high potential (His) (axial
            ligands) #status predicted
SUMMARY      #length 215 #molecular-weight 24166 #checksum 9582

```

Query Match 5.5%; Score 106; DB 1; Length 215;  
Best Local Similarity 22.7%; Pred. No. 5.58e-01;  
Matches 15; Conservative 21; Mismatches 24; Indels 6; Gaps 6;

Db 57 YYPRTVDTAFASVQYIMTEVNFGLI-R-SV-HRWSASMVLMLHVFVYLTTGGFKP 113  
| : :: :: |::: | : | : | : :: | : | : :: |  
Qy 132 YIQMSIDSNISLVHYIVASAQV-WMITRYDLYHNFRPAVL-LLMFLSVYKAFVMTFVHL 189

Db 114 REL-TW 118  
| : |

Qy 190 CSLGSW 195

```

RESULT 13
ENTRY      B69031      #type complete
TITLE      conserved hypothetical protein MTH1229 - Methanobacterium
            thermoautotrophicum (strain Delta H)
ORGANISM    #formal_name Methanobacterium thermoautotrophicum
DATE        05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change
            13-Sep-1998
ACCESSIONS  B69031
REFERENCE   A69000
            #authors    Smith, D.R.; Doucette-Stamm, L.A.; Delouhery, C.; Lee, H.;

```

Dubois, J.; Aldredge, T.; Bashirzadeh, R.; Blakely, D.; Cook, R.; Gilbert, K.; Harrison, D.; Hoang, L.; Keagle, P.; Lum, W.; Pothier, B.; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.; Caruso, A.; Bush, D.; Safer, H.; Patwell, D.; Prabhakar, S.; McDougall, S.; Shimer, G.; Goyal, A.; Pietrokovski, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.

#journal J. Bacteriol. (1997) 179:7135-7155  
#title Complete genome sequence of *Methanobacterium thermoautotrophicum* Delta H: functional analysis and comparative genomics.

#cross-references MUID:98037514

#accession B69031

##status preliminary; nucleic acid sequence not shown;  
translation not shown

##molecule\_type DNA

##residues 1-212 ##label MTH

##cross-references GB:AE000890; GB:AE000666; NID:g2622331; PID:g2622340

##experimental\_source strain Delta H

## GENETICS

#gene MTH1229

CLASSIFICATION #superfamily *Methanococcus jannaschii* conserved hypothetical protein MJ0805

SUMMARY #length 212 #molecular-weight 23956 #checksum 2161

Query Match 5.4%; Score 105; DB 2; Length 212;  
Best Local Similarity 31.5%; Pred. No. 7.28e-01;  
Matches 23; Conservative 17; Mismatches 27; Indels 6; Gaps 6;

Db 44 AWALSELGKSSIIITWIGPMDVKNIVDYEDVELEIRSEEVLFHIVEHFE-QPSSLRL 101

Qy 105 GWATAEL-IMSRCIPLWVGARGI-EF-DWKYIQMSIDSNISLVHYIVASAQVWMITRYDL 161

Db 102 AYHRQRIVLMLLM 114

Qy 162 -YHNERPAVLLM 173

## RESULT 14

ENTRY CBLV6 #type complete

TITLE plastoquinol-plastocyanin reductase (EC 1.10.99.1)  
cytochrome b6 - liverwort (*Marchantia polymorpha*)  
chloroplast

ORGANISM #formal\_name chloroplast *Marchantia polymorpha*

DATE 30-Jun-1987 #sequence\_revision 30-Jun-1987 #text\_change  
04-Oct-1996

ACCESSIONS S01552; S02432; A00163

REFERENCE S01529

#authors Fukuzawa, H.; Kohchi, T.; Sano, T.; Shirai, H.; Umesono, K.;  
Inokuchi, H.; Ozeki, H.; Ohyama, K.

#journal J. Mol. Biol. (1988) 203:333-351

#title Structure and organization of *Marchantia polymorpha*  
chloroplast genome. III. Gene organization of the large  
single copy region from *rbcl* to *trnI*(CAU).

#cross-references MUID:89068687

#accession S01552

##molecule\_type DNA

##residues 1-215 ##label FUK

##cross-references EMBL:X04465

REFERENCE A38014

#authors Ohyama, K.; Fukuzawa, H.; Kohchi, T.; Shirai, H.; Sano, T.;  
Sano, S.; Umesono, K.; Shiki, Y.; Takeuchi, M.; Chang, Z.;  
Aota, S.; Inokuchi, H.; Ozeki, H.

#journal Nature (1986) 322:572-574

#title Chloroplast gene organization deduced from complete sequence  
of liverwort *Marchantia polymorpha* chloroplast DNA.

#contents annotation; gene organization, sites, features

REFERENCE S02432

#authors Fukuzawa, H.; Yoshida, T.; Kohchi, T.; Okumura, T.; Sawano,  
Y.; Ohyama, K.

#journal FEBS Lett. (1987) 220:61-66

#title Splicing of group II introns in mRNAs coding for cytochrome

b6 and subunit IV in the liverwort *Marchantia polymorpha*  
chloroplast genome. Exon specifying a region coding for two  
genes with the spacer region.

#accession S02432

##status not compared with conceptual translation

##molecule\_type mRNA

##residues 1-8 ##label FUV

## GENETICS

#gene petB

#genome chloroplast

#introns 2/3

CLASSIFICATION #superfamily cytochrome b6; cytochrome b6 homology

KEYWORDS chloroplast; chromoprotein; electron transfer; heme; iron;  
oxidoreductase; photosynthesis; thylakoid

## FEATURE

15-215 #domain cytochrome b6 homology #label CB6\

86,187 #binding\_site heme iron, low potential (His) (axial  
ligands) #status predicted\

100,202 #binding\_site heme iron, high potential (His) (axial  
ligands) #status predicted

SUMMARY #length 215 #molecular-weight 24307 #checksum 8453

Query Match 5.4%; Score 105; DB 1; Length 215;

Best Local Similarity 22.7%; Pred. No. 7.28e-01;

Matches 15; Conservative 21; Mismatches 24; Indels 6; Gaps 6;

Db 57 YYPRTVTEAFSSVQYIMTEVNFGLI-R-SV-HRNSASMMVLMILHIFRVYLTGGFKKP 113

Qy 132 YIQMSIDSNISLVHYIVASAQV-WMITRYDLYHNERPAVL-LLMFLSVYKAFVMTFVHL 189

Db 114 REL-TW 118

Qy 190 CSLGSW 195

## RESULT 15

ENTRY S73262 #type complete

TITLE cytochrome b6 - red alga (*Porphyrha purpurea*) chloroplast

ORGANISM #formal\_name chloroplast *Porphyrha purpurea*

DATE 19-Mar-1997 #sequence\_revision 09-May-1997 #text\_change  
24-Apr-1998

ACCESSIONS S73262

REFERENCE S73108

#authors Reith, M.; Munholland, J.

#journal Plant Mol. Biol. Rep. (1995) 13:333-335

#title Complete nucleotide sequence of the *Porphyrha purpurea*  
chloroplast genome.

#accession S73262

##status preliminary; nucleic acid sequence not shown;  
translation not shown

##molecule\_type DNA

##residues 1-215 ##label REI

##cross-references EMBL:U38804; NID:g1276652; PID:g1276807

##note the nucleotide sequence was submitted to the EMBL Data  
Library, October 1995

## GENETICS

#gene petB

#genome chloroplast

CLASSIFICATION #superfamily cytochrome b6; cytochrome b6 homology

KEYWORDS chloroplast; thylakoid

## FEATURE

15-215 #domain cytochrome b6 homology #label CB6

SUMMARY #length 215 #molecular-weight 24224 #checksum 9699

Query Match 5.4%; Score 104; DB 2; Length 215;

Best Local Similarity 22.7%; Pred. No. 9.48e-01;

Matches 15; Conservative 20; Mismatches 25; Indels 6; Gaps 6;

Db 57 YYPRTVAEFTSVEYIMTVNFGLI-R-SI-HRNSASMMVLMILHIFRVYLTGGFKKP 113

Qy 132 YIQMSIDSNISLVHYIVASAQV-WMITRYDLYHNERPAVL-LLMFLSVYKAFVMTFVHL 189

Db 114 REL-TW 118

Sun Jan 24 11:29:02 1999

US-09-162-597-3.rpr

Page 7

Qy 190 CSLGSW 195

Search completed: Fri Jan 22 18:00:22 1999  
Job time : 19 secs.

\*\*\*\*\*

MISOL (TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Fri Jan 22 18:03:07 1999; MasPar time 3.58 Seconds  
597.255 Million cell updates/sec

Tabular output not generated.

Title: >US-09-162-597-3  
Description: (1-245) from US09162597.pep  
Perfett Score: 1940  
Sequence: 1 MTLFHFNGCFALAYFPYFIT.....GLVSQTLMYLFPASLQVLVK 245

Scoring table: PAM 150  
Gap 11

Searched: 92929 seqs, 8738560 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: a-issued  
1:5\_COMB 2:PCT9\_COMB 3:backfiles1

Statistics: Mean 31.166; Variance 148.701; scale 0.210

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result	No.	Score	Match	Length	DB	ID	Description	Pred. No.
1	94	4.8	772	2	PCT-US95-0	Sequence 2, Applicatio	1.66e+01	
2	94	4.8	772	1	US-08-258-	Sequence 2, Applicatio	1.66e+01	
3	89	4.6	349	1	US-08-118-	Sequence 71, Applicati	3.57e+01	
4	89	4.6	349	2	PCT-US93-0	Sequence 71, Applicati	3.57e+01	
5	87	4.5	307	1	US-08-948-	Sequence 3, Applicatio	4.82e+01	
6	87	4.5	417	1	US-08-553-	Sequence 2, Applicatio	4.82e+01	
7	85	4.4	201	3	5489533-3	Patent No. 5489533.	6.50e+01	
8	85	4.4	252	2	PCT-US95-1	Sequence 103, Applicat	6.50e+01	
9	85	4.4	274	3	5489533-2	Patent No. 5489533.	6.50e+01	
10	85	4.4	274	3	5512660-2	Patent No. 5512660.	6.50e+01	
11	84	4.3	316	1	US-08-118-	Sequence 46, Applicati	7.54e+01	
12	84	4.3	316	2	PCT-US93-0	Sequence 46, Applicati	7.54e+01	
13	83	4.3	346	2	PCT-US96-1	Sequence 2, Applicatio	8.75e+01	
14	83	4.3	363	1	US-08-148-	Sequence 2, Applicatio	8.75e+01	
15	83	4.3	363	1	US-08-148-	Sequence 3, Applicatio	8.75e+01	
16	83	4.3	380	2	PCT-US93-1	Sequence 40, Applicati	8.75e+01	
17	83	4.3	380	1	US-08-153-	Sequence 40, Applicati	8.75e+01	
18	84	4.3	494	2	PCT-US94-0	Sequence 4, Applicatio	7.54e+01	
19	84	4.3	494	1	US-08-464-	Sequence 4, Applicatio	7.54e+01	
20	84	4.3	2368	1	US-08-198-	Sequence 15, Applicati	7.54e+01	
21	82	4.2	225	3	5436139-4	Patent No. 5436139.	1.01e+02	
22	82	4.2	226	3	5436139-5	Patent No. 5436139.	1.01e+02	
23	82	4.2	226	3	5196194-21	Patent No. 5196194.	1.01e+02	

24	82	4.2	226	2	PCT-US96-1	Sequence 14, Applicati	1.01e+02
25	82	4.2	226	3	5198348-1	Patent No. 5198348.	1.01e+02
26	82	4.2	226	1	US-08-378-	Sequence 3, Applicatio	1.01e+02
27	82	4.2	236	1	US-08-378-	Sequence 1, Applicatio	1.01e+02
28	82	4.2	277	3	5164485-2	Patent No. 5164485.	1.01e+02
29	82	4.2	279	1	US-07-688-	Sequence 30, Applicati	1.01e+02
30	82	4.2	279	2	PCT-US91-0	Sequence 29, Applicati	1.01e+02
31	82	4.2	281	1	US-08-458-	Sequence 214, Applicat	1.01e+02
32	82	4.2	281	1	US-08-105-	Sequence 214, Applicat	1.01e+02
33	82	4.2	351	2	PCT-US96-1	Sequence 4, Applicatio	1.01e+02
34	82	4.2	361	2	PCT-US93-0	Sequence 4, Applicatio	1.01e+02
35	82	4.2	389	1	US-08-709-	Sequence 216, Applicat	1.01e+02
36	82	4.2	389	1	US-08-709-	Sequence 219, Applicat	1.01e+02
37	82	4.2	389	1	US-08-105-	Sequence 219, Applicat	1.01e+02
38	82	4.2	389	1	US-08-458-	Sequence 216, Applicat	1.01e+02
39	82	4.2	389	1	US-08-458-	Sequence 219, Applicat	1.01e+02
40	82	4.2	389	1	US-08-105-	Sequence 216, Applicat	1.01e+02
41	82	4.2	395	3	5196194-18	Patent No. 5196194.	1.01e+02
42	82	4.2	397	2	PCT-US96-1	Sequence 6, Applicati	1.01e+02
43	82	4.2	502	1	US-08-496-	Sequence 4, Applicatio	1.01e+02
44	82	4.2	502	1	US-08-466-	Sequence 10, Applicati	1.01e+02
45	82	4.2	995	2	PCT-US95-0	Sequence 14, Applicati	1.01e+02

## ALIGNMENTS

RESULT 1

ID PCT-US95-07391A-2 STANDARD; PRT; 772 AA.

XX

AC xxxxxx

XX

DT

XX

DE Sequence 2, Application PC/TUS9507391A

XX

CC Sequence 2, Application PC/TUS9507391A

CC GENERAL INFORMATION:

CC APPLICANT: IBEX TECHNOLOGIES and

CC APPLICANT: ZIMMERMANN, Joseph

CC TITLE OF INVENTION: Nucleic Acid Sequences And Expression

CC TITLE OF INVENTION: Systems For Heparinase II And Heparinase III Derived Fr

CC TITLE OF INVENTION: Flavobacterium heparinum

CC NUMBER OF SEQUENCES: 26

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Hale and Dorr

CC STREET: 1455 Pennsylvania Avenue, N.W.

CC CITY: Washington, D.C.

CC COUNTRY: U.S.A.

CC ZIP: 20004

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Floppy disk

CC COMPUTER: IBM PC compatible

CC OPERATING SYSTEM: PC-DOS/MS-DOS

CC SOFTWARE: PatentIn Release #1.0, Version #1.25

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: PCT/US95/07391A

CC FILING DATE: 09-JUNE-1995

CC CLASSIFICATION:

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: 08/258,639

CC FILING DATE: 10 JUNE 1994

CC ATTORNEY/AGENT INFORMATION:

CC NAME: BAKER, Hollie L.

CC REGISTRATION NUMBER: 31,321

CC REFERENCE/DOCKET NUMBER: 104385.116PCT

CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: (202)942-8400

CC TELEFAX: (202)942-8484

CC INFORMATION FOR SEQ ID NO: 2:

CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 772 amino acids

CC TYPE: amino acid

CC TOPOLOGY: linear

Qy 184 ETFVHLC SLG SWARLDARA-VVTGLLAL 210



CC Sequence 2, Application US/08553703A  
CC Patent No. 5795767  
CC GENERAL INFORMATION:  
CC APPLICANT: MARU, ISAFUMI  
CC APPLICANT: OHTA, YASUHIRO  
CC APPLICANT: TSUKADA, YOJI  
CC TITLE OF INVENTION: EPIMERASE  
CC NUMBER OF SEQUENCES: 9  
CC CORRESPONDENCE ADDRESS:

RESULT 8  
ID PCT-US95-14442A-103 STANDARD; PRT; 252 AA.  
XX  
AC xxxxxx  
XX  
DT  
XX  
DE Sequence 103, Application PC/TUS9514442A  
XX  
CC Sequence 103, Application PC/TUS9514442A  
CC GENERAL INFORMATION:  
CC APPLICANT: Grieve, Robert B.  
CC APPLICANT: Rushlow, Keith E.  
CC APPLICANT: Hunter, Shirley Wu  
CC APPLICANT: Frank, Glenn R.  
CC APPLICANT: Heath, Andrew W.  
CC APPLICANT: Yamaka, Miles Yamanaka  
CC APPLICANT: Arfsten, Ann  
CC APPLICANT: Dale, Beverly  
CC APPLICANT: Stiegler, Gary  
CC TITLE OF INVENTION: USE OF PROTEASE INHIBITORS AND  
CC TITLE OF INVENTION: PROTEASE VACCINES TO PROTECT ANIMALS FROM FLEA  
CC TITLE OF INVENTION: INFESTATION, AND FLEA PROTEASE PROTEINS, NUCLEIC ACID  
CC TITLE OF INVENTION: MOLECULES, AND USES THEREOF  
CC NUMBER OF SEQUENCES: 119  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Sheridan Ross & McIntosh  
CC STREET: 1700 Lincoln Street, Suite 3500  
CC CITY: Denver  
CC STATE: Colorado  
CC COUNTRY: USA  
CC ZIP: 80203  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: PCT/US95/14442A  
CC FILING DATE:  
CC CLASSIFICATION:  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Gary J. Connell  
CC REGISTRATION NUMBER: 32,020  
CC REFERENCE/DOCKET NUMBER:  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (303) 863-9700  
CC TELEFAX: (303) 863-0223  
CC INFORMATION FOR SEO ID NO: 103:

CC NAME: Townsend, Kevin G.

RESULT 12  
ID PCT-US93-08528-46 STANDARD; PRT; 316 AA.  
XX  
AC xxxxxx  
XX  
XX  
DT  
XX  
DE Sequence 46, Application PC/TUS9308528  
XX  
CC Sequence 46, Application PC/TUS9308528  
CC GENERAL INFORMATION:  
CC APPLICANT: New York University  
CC TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN  
CC TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF  
CC NUMBER OF SEQUENCES: 348  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: BROWDY AND NEIMARK  
CC STREET: 419 Seventh Street, N.W., Suite 300  
CC CITY: Washington  
CC STATE: D.C.  
CC COUNTRY: USA  
CC ZIP: 20004  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: PCT/US93/08528  
CC FILING DATE: 09-SEP-1993  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 07/943,236  
CC FILING DATE: 10-SEP-1992  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Townsend, Kevin G.  
CC REGISTRATION NUMBER: 34,033  
CC REFERENCE/DOCKET NUMBER: MURPHY-2 PCT  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 202-628-5197  
CC TELEFAX: 202-737-3528  
CC TELEX: 248633  
CC INFORMATION FOR SEQ ID NO: 46:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 316 amino acids  
CC TYPE: amino acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: peptide  
SQ SEQUENCE 316 AA; 36040 MW; 553094 CN;

```

RESULT 13
ID PCT-US96-10602-2 STANDARD; PRT; 346 AA.
XX
AC XXXXXX
XX
XX
DT
XX
DE Sequence 2, Application PC/TUS9610602
XX
CC Sequence 2, Application PC/TUS9610602
CC GENERAL INFORMATION:
CC APPLICANT: The General Hospital Corporation
CC TITLE OF INVENTION: INHIBITION OF HEPATITIS B REPLICATION
CC NUMBER OF SEQUENCES: 14
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Fish & Richardson P.C.
CC STREET: 225 Franklin Street
CC CITY: Boston
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 02110-2804
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US96/10602
CC FILING DATE:
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 60/017,814
CC FILING DATE: 20-JUN-1995
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Clark, Paul T.
CC REGISTRATION NUMBER: 30,162
CC REFERENCE/DOCKET NUMBER: 00786/282001
CC TELECOMMUNICATION INFORMATION: ...
CC TELEPHONE: 617/542-5070
CC TELEFAX: 617/542-8906
CC TELEX: 200154
CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 346 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
SQ SEQUENCE 346 AA; 40128 MW; 678897 CN;

```

RESULT 14  
ID US-08-148-209A-2 STANDARD; PRT; 363 AA.  
XX  
AC XXXXXX



\*\*\*\*\*

RELEASE (TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Fri Jan 22 17:58:21 1999; MasPar time 8.04 Seconds  
492.876 Million cell updates/sec

Tabular output not generated.

Title: >US-09-162-597-3  
Description: (1-245) from US09162597.pep  
Perfect Score: 1940  
Sequence: 1 MTLFHFNGCFALAYFPYFIT.....GLVSQTLMYLFPASLQVLVK 245

Scoring table: PAM 150  
Gap 11

Searched: 131922 seqs, 16180660 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: a-geneseq32  
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
14:part14 15:part15 16:part16 17:part17 18:part18  
19:part19 20:part20 21:part21 22:part22 23:part23  
24:part24 25:part25 26:part26 27:part27 28:part28  
29:part29

Statistics: Mean 33.119; Variance 147.508; scale 0.225

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	No.	Score	Match	Length	DB	ID	Description	Pred. No.
1	99	5.1	237	22	W20142	H. pylori inner membr	1.14e+01	
2	98	5.1	422	21	R97722	Mouse inositol polyph	1.34e+01	
3	98	5.1	497	22	W20821	H. pylori cell envelo	1.34e+01	
4	95	4.9	1353	19	R99251	Murine adenylate cycl	2.17e+01	
5	94	4.8	772	15	R89264	Heparinase-II.	2.55e+01	
6	94	4.8	843	13	R67760	Lys-aminopeptidase Pe	2.55e+01	
7	91	4.7	359	22	W20927	H. pylori surface or	4.11e+01	
8	92	4.7	521	21	W14445	CarA gene product.	3.51e+01	
9	89	4.6	349	19	W02722	G-protein coupled mou	5.64e+01	
10	89	4.6	349	16	R48750	G-protein coupled mou	5.64e+01	
11	87	4.5	174	3	R21412	NADH dehydrogenase 6.	7.70e+01	
12	87	4.5	293	21	W14481	Ranp-1.	7.70e+01	
13	87	4.5	417	16	R79929	Porcine acylglucosani	7.70e+01	
14	88	4.5	558	2	P70429	Murine neuroleukin.	6.59e+01	
15	85	4.4	152	29	W48271	Rat ninjurin 1.	1.05e+02	
16	85	4.4	252	20	W01210	Serine protease PfSP1	1.05e+02	
17	85	4.4	274	1	R06836	Intercellular adhesio	1.05e+02	
18	85	4.4	1477	22	W10424	Saccharomyces cerevis	1.05e+02	

19	85	4.4	1477	13	R67691	S. cerevisiae scaur2R	1.05e+02
20	86	4.4	1989	18	R99640	Peripheral nervous sy	8.99e+01
21	86	4.4	4473	20	R97244	Virulence gene cluste	8.99e+01
22	84	4.3	236	1	P81992	Hepatitis B viral sur	1.22e+02
23	84	4.3	316	19	W02697	G-protein coupled rat	1.22e+02
24	84	4.3	316	16	R48725	G-protein coupled rat	1.22e+02
25	83	4.3	346	21	W09046	WHV core-surface fusi	1.42e+02
26	83	4.3	363	13	R66934	Mouse AT2 receptor.	1.42e+02
27	83	4.3	380	11	R53750	Seven transmembrane r	1.42e+02
28	83	4.3	391	29	W38453	Wild-type human homol	1.42e+02
29	84	4.3	422	22	W20249	H. pylori transembra	1.22e+02
30	84	4.3	426	22	W20934	H. pylori surface or	1.22e+02
31	84	4.3	494	28	W42996	Putative mature potas	1.22e+02
32	84	4.3	494	16	R90765	Human K+ channel 2 ma	1.22e+02
33	84	4.3	1233	29	W34536	Nudaurelia beta virus	1.22e+02
34	84	4.3	1233	29	W41935	Nudaurelia beta-like	1.22e+02
35	84	4.3	2368	26	W26663	Yeast checkpoint cont	1.22e+02
36	84	4.3	2510	6	R29527	HCV antigen T7N1-30.	1.22e+02
37	84	4.3	3011	4	R22154	NANBV Hutch c59 isola	1.22e+02
38	83	4.3	3085	23	W19701	ATM mutant G9170C.	1.42e+02
39	82	4.2	226	7	R33252	HBsAg encoded by pGPD	1.66e+02
40	82	4.2	250	2	R11496	RP142/HBsAg.	1.66e+02
41	82	4.2	351	21	W09047	Plasmid pHBV DN inser	1.66e+02
42	82	4.2	558	2	P70430	Human neuroleukin.	1.66e+02
43	82	4.2	658	27	W27666	Streptococcus pneumon	1.66e+02
44	82	4.2	724	27	W36797	Novel human gene, des	1.66e+02
45	82	4.2	898	21	W14777	Granulosis virus infe	1.66e+02

#### ALIGNMENTS

RESULT 1

ID W20142 standard; protein; 237 AA.

AC W20142;

DT 08-JUL-1997 (first entry)

DE H. pylori inner membrane protein, 14455461.aa.

KW Cytoplasmic; vaccine; prevention; treatment; infection; identification;

KW binding compound; bacterium; life cycle; activator; bacteria; inhibitor;

KW duodenal ulcer disease; chronic gastritis; diagnosis; envelope.

OS Helicobacter pylori.

FH Key Location/Qualifiers

FT misc\_difference 3

FT /note= "encoded by YCC"

FT misc\_difference 130

FT /note= "encoded by RCG"

FT misc\_difference 183

FT /note= "encoded by GSY"

PN W09640893-A1.

PD 19-DEC-1996.

PF 06-JUN-1996; U09122.

PR 07-JUN-1995; US-487032.

PR 01-APR-1996; US-630405.

PA (ASTR ) ASTRA AB.

PI Berglindh OT, Smith D, Mellgaard BL;

DR WPI; 97-052306/05.

DR N-PSDB; T67385.

PT Helicobacter pylori nucleic acid sequences and related

PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori

PT infection, and to detect Helicobacter

PS Claim 56; Pages 365; 1481pp; English.

CC The present sequence is a Helicobacter pylori inner membrane protein.

CC The protein may be used in a vaccine to prevent or treat H. pylori

CC infection or to identify H. pylori polypeptide binding compounds,

CC useful as potential H. pylori life cycle activators or inhibitors.

CC The genomic sequence of H. pylori (ATCC 55679) was determined from

CC overlapping contigs generated by mechanically shearing the bacterial

CC DNA. The sequences were analysed for ORF of at least 180 nucleotides,

CC and the predicted coding regions defined by computer evaluation. To

CC identify likely H. pylori antigens for vaccine development, the amino

CC acid sequences predicted from various ORF were analysed for significant

CC homology to other known or exported membrane proteins. Having identified

CC and determined the sequences of interest, particular regions can be

CC isolated from H. pylori by PCR amplification for recombinant polypeptide

Key	Location/Qualifiers
modified_site	27
	/label= C-kinase
	/note= "putative protein kinase C phosphorylation site"
modified_site	84
	/label= CAM_kinase-II
	/note= "putative CAM kinase II phosphorylation site"
modified_site	85
	/label= A-kinase
	/note= "putative protein kinase A phosphorylation site"
modified site	94

FT /label= C-kinase  
 FT /note= "putative protein kinase C phosphorylation  
 FT site"  
 FT modified\_site 187  
 FT /label= C-kinase  
 FT /note= "putative protein kinase C phosphorylation  
 FT site"  
 FT modified\_site 200  
 FT /label= Cam\_kinase-II  
 FT /note= "putative CAM kinase II phosphorylation  
 FT site"  
 FT modified\_site 201  
 FT /label= A-kinase  
 FT /note= "putative protein kinase A phosphorylation  
 FT site"  
 FT modified\_site 261  
 FT /label= Casein\_kinase-II  
 FT /note= "putative casein kinase II phosphorylation  
 FT site"  
 FT modified\_site 266  
 FT /label= C-kinase  
 FT /note= "putative protein kinase C phosphorylation  
 FT site"  
 FT modified\_site 303  
 FT /label= CAM\_kinase-II  
 FT /note= "putative CAM kinase II phosphorylation  
 FT site"  
 FT modified\_site 304  
 FT /label= A-kinase  
 FT /note= "putative protein kinase A phosphorylation  
 FT site"  
 FT modified\_site 356  
 FT /label= C-kinase  
 FT /note= "putative protein kinase C phosphorylation  
 FT site"  
 FT modified\_site 358  
 FT /label= CAM\_kinase-II  
 FT /note= "putative CAM kinase II phosphorylation  
 FT site"  
 FT modified\_site 370  
 FT /label= C-kinase  
 FT /note= "putative protein kinase C phosphorylation  
 FT site"  
 FT modified\_site 432  
 FT /label= Casein\_kinase-II  
 FT /note= "putative casein kinase II phosphorylation  
 FT site"  
 FT modified\_site 436  
 FT /label= C-kinase  
 FT /note= "putative protein kinase C phosphorylation  
 FT site"  
 FT domain 503..610  
 FT /label= Calcineurin\_binding\_domain  
 FT modified\_site 581  
 FT /label= Casein\_kinase-II  
 FT /note= "putative casein kinase II phosphorylation  
 FT site"  
 FT modified\_site 588  
 FT /label= Casein\_kinase-II  
 FT /note= "putative casein kinase II phosphorylation  
 FT site"  
 FT modified\_site 596  
 FT /label= CAM\_kinase-II  
 FT /note= "putative CAM kinase II phosphorylation  
 FT site"  
 FT modified\_site 597  
 FT /label= A-kinase  
 FT /note= "putative protein kinase A phosphorylation  
 FT site"  
 FT modified\_site 619  
 FT /label= C-kinase  
 FT /note= "putative protein kinase C phosphorylation  
 FT site"

FT modified\_site 690  
 FT /label= Casein\_kinase-II  
 FT /note= "putative casein kinase II phosphorylation  
 FT site"  
 FT modified\_site 719  
 FT /label= CAM\_kinase-II  
 FT /note= "putative CAM kinase II phosphorylation  
 FT site"  
 FT modified\_site 732  
 FT /label= C-kinase  
 FT /note= "putative protein kinase C phosphorylation  
 FT site"  
 FT modified\_site 761  
 FT /label= C-kinase  
 FT /note= "putative protein kinase C phosphorylation  
 FT site"  
 FT modified\_site 763  
 FT /label= A-kinase  
 FT /note= "putative protein kinase A phosphorylation  
 FT site"  
 FT modified\_site 765  
 FT /label= Casein\_kinase-II  
 FT /note= "putative casein kinase II phosphorylation  
 FT site"  
 FT modified\_site 837  
 FT /label= C-kinase  
 FT /note= "putative protein kinase C phosphorylation  
 FT site"  
 FT modified\_site 885  
 FT /label= Casein\_kinase-II  
 FT /note= "putative casein kinase II phosphorylation  
 FT site"  
 FT modified\_site 919  
 FT /label= A-kinase  
 FT /note= "putative protein kinase A phosphorylation  
 FT site"  
 FT modified\_site 971  
 FT /label= CAM\_kinase-II  
 FT /note= "putative CAM kinase II phosphorylation  
 FT site"  
 FT modified\_site 972  
 FT /label= A-kinase  
 FT /note= "putative protein kinase A phosphorylation  
 FT site"  
 FT modified\_site 972  
 FT /label= C-kinase  
 FT /note= "putative protein kinase C phosphorylation  
 FT site"  
 FT modified\_site 1002  
 FT /label= C-kinase  
 FT /note= "putative protein kinase C phosphorylation  
 FT site"  
 FT modified\_site 1019  
 FT /label= C-kinase  
 FT /note= "putative protein kinase C phosphorylation  
 FT site"  
 FT modified\_site 1002  
 FT /label= C-kinase  
 FT /note= "putative protein kinase C phosphorylation  
 FT site"  
 FT modified\_site 1105  
 FT /label= C-kinase  
 FT /note= "putative protein kinase C phosphorylation  
 FT site"  
 FT modified\_site 1194  
 FT /label= CAM\_kinase-II  
 FT /note= "putative CAM kinase II phosphorylation  
 FT site"  
 FT modified\_site 1199  
 FT /label= Casein\_kinase-II  
 FT /note= "putative casein kinase II phosphorylation  
 FT site"  
 FT modified\_site 1211



FT /label= C-kinase  
FT /note= "putative protein kinase C phosphorylation  
FT site"  
FT modified\_site 1212  
FT /label= CAM\_kinase-II  
FT /note= "putative CAM kinase II phosphorylation  
FT site"  
FT modified\_site 1213  
FT /label= A-kinase  
FT /note= "putative protein kinase A phosphorylation  
FT site"  
FT modified\_site 1270  
FT /label= CAM\_kinase-II  
FT /note= "putative CAM kinase II phosphorylation  
FT site"  
FT modified\_site 1310  
FT /label= Casein\_kinase-II  
FT /note= "putative casein kinase II phosphorylation  
FT site"  
FT modified\_site 1310  
FT /label= C-kinase  
FT /note= "putative protein kinase C phosphorylation  
FT site"  
PN W09625502-A1.  
PD 22-AUG-1996.  
PF 14-FEB-1996; G00312.  
PR 14-FEB-1995; GB-002806.  
PR 11-AUG-1995; GB-016528.  
PA (MEDI-) MEDICAL RES COUNCIL.  
PI Antoni F, Paterson JM;  
DR WPI; 96-393403/39.  
DR N-PSDB; T35209.

Note: remainder of annotations omitted.

Query Match 4.9%; Score 95; DB 19; Length 1353;  
Best Local Similarity 26.8%; Pred. No. 2.17e+01;  
Matches 15; Conservative 16; Mismatches 23; Indels 2; Gaps 2;

Db 887 efetnihvtmftgsavlvavvhyhcnfcqlsswmrslatiagallllhis-lcq 941  
:| : : || : : : : || : : : || : : : :  
Qy 164 NFRPAVLLMFLSVYKAFVMTFVHLCSLGSWALDARAVV-TGLLALKHFGPVC R 218

## RESULT 5

ID R89264 standard; Protein; 772 AA.  
AC R89264;  
DT 07-APR-1996 (first entry)  
DE Heparinase-II.  
KW Heparinase-II; heparin degradation; heparan sulphate degradation;  
KW Escherichia coli; polymerase chain reaction; PCR; primer; toxicity;  
KW cloning; vector; ribosome binding site; haemostatic;  
KW blood-clotting; antibody; affinity chromatography.  
OS Flavobacterium heparinum.  
FH Key Location/Qualifiers  
FT peptide 1..25  
FT /note= "Signal peptide"  
FT peptide 85..96  
FT /note= "Peptide 2B (R89267)"  
FT peptide 222..233  
FT /note= "Peptide 2A (R89266)"  
FT peptide 500..526  
FT /note= "Peptide 2C (R89268)"

PN W09534635-A1.  
PD 21-DEC-1995.  
PF 09-JUN-1995; U07391.  
PR 10-JUN-1994; US-258639.  
PA (IBEX-) IBEX TECHNOLOGIES.  
PA (ZIMM-) ZIMMERMANN J.  
PI Bennett C, Blain F, Gu K, Musil R, Su H, Zimmermann J;  
DR WPI; 96-097381/10.  
DR N-PSDB; Q99226.  
PT Nucleic acids encoding Flavobacterium heparinum heparinase II and

PT III - for degrading heparin and heparan sulphate, also related host  
PT cells, proteins and antibodies, useful in heparinase purificn. ....  
PS Claim 14; Fig 5; 75pp; English.  
CC The sequence corresponds to Flavobacterium heparinum heparinase-II,  
CC which degrades heparin and heparan sulphate. The sequence is  
CC isolated from a gene library in phage lambda-DASH-II in Escherichia  
CC coli by polymerase chain reaction amplification using e.g. primers  
CC 2-1 to 2-4 (Q99228-Q99231), which are designed based on the  
CC sequences of peptides 2A and 2B. Toxicity of the gene and natural  
CC selection of the host against clones with the entire sequence has  
CC been circumvented by cloning sections of the gene separately in a  
CC vector with a modified ribosome binding site, which increases  
CC expression levels. The heparinase-II may be used to neutralise  
CC anticoagulant activity. Antibodies against the protein may be used  
CC to differentiate between native and recombinant enzymes, and when  
CC immobilised they may be used for heparinase purification by  
CC affinity chromatography.  
SQ Sequence 772 AA;

Query Match 4.8%; Score 94; DB 15; Length 772;  
Best Local Similarity 26.1%; Pred. No. 2.55e+01;  
Matches 23; Conservative 25; Mismatches 35; Indels 5; Gaps 3;

Db 94 dfrfyfngkgltrvrmalnlylmtkdpkgreatsiidltetatkpagdisrgiglf 153  
:| : : || : : : : || : : : || : : : :  
Qy 66 DFIEFMKASVDV-ADLIGLNLVMSRNA--GKGEYKIMVAALGWAT--AELIMSRCIPLW 120  
Db 154 mvtgaivdywcydqlkpeektrfvkafv 181  
:| : : || : : : : || : : : || : : : :  
Qy 121 VGARGIEFDWKYIQMSIDSNISLVHYIV 148

## RESULT 6

ID R67760 standard; Protein; 843 AA.  
AC R67760;  
DT 01-AUG-1995 (first entry)  
DE Lys-aminopeptidase PepN.  
KW Lys-aminopeptidase; PepN; fermented food; cheese.  
OS Lactobacillus delbrueckii subsp. lactis.  
PN EP-633316-A.  
PD 11-JAN-1995.  
PF 30-JUN-1994; 401497.  
PR 01-JUL-1993; GB-013586.  
PA (EECE-) EEC EURO ECONOMIC COMMUNITY.  
PI Klein JR, Plapp R;  
DR WPI; 95-038513/06.  
DR N-PSDB; Q79913.  
PT Purified Lys-aminopeptidase PepN enzyme and PepN gene - useful  
PT for prepn. of fermented foodstuff, esp. cheese.  
PS Claim 3; Page 17-21; 41pp; English.  
CC A new Lys-aminopeptidase, PepN, was isolated from L. delbrueckii  
CC subsp. lactis WS87 (DSM 7290) and had the sequence given in  
CC R67760. The pepN gene (Q79913) was isolated from a library  
CC of DSM 7290 chromosomal DNA by screening for peptolytic activity  
CC in Escherichia coli ER1562 transformants. The isolated gene  
CC is used for recombinant PepN production.  
SQ Sequence 843 AA;

Query Match 4.8%; Score 94; DB 13; Length 843;  
Best Local Similarity 21.1%; Pred. No. 2.55e+01;  
Matches 8; Conservative 15; Mismatches 14; Indels 1; Gaps 1;

Db 761 ewawlektvggdmefatitvisrvfktkerydeynaf 798  
:| : : || : : : : || : : : || : : : :  
Qy 129 DWKYIQMSIDSNISLVHYIVASQVWMIT-RYDLYHNF 165

## RESULT 7

ID W20927 standard; protein; 359 AA.  
AC W20927;  
DT 21-JUL-1997 (first entry)  
DE H. pylori surface or membrane protein, 16ael0508orf3.  
KW Cytoplasmic; vaccine; prevention; treatment; infection; identification;

RESULT 8

ID W14445 standard; Protein; 521 AA.

AC W14445;

DT 17-MAY-1997 (first entry)

DE CarA gene product.

KW carbapenem; carR; carA; carB; carC; carD; carE; carF; carG; carH;

KW biosynthesis; antibiotic; beta-lactamase inhibitor; regulatory.

OS *Erwinia carotovora*.

PN W09532294-A1.

PD 30-NOV-1995.

PR 18-MAY-1995; G01125.

PR 20-MAY-1994; GB-010142.

PA (UYNO-) UNIV NOTTINGHAM. PA (UYWA-) UNIV WARWICK.

PI. Bycroft BW, Cox ARJ, Holden MTG, McGowan SJ, Porter LE;

PI Salmond GPC, Sebahia M, Stewart GSA, Williams P;

DR WPI; 96-020587/02.

DR N-PSDB; T09700.

PT New genes encoding enzymes involved in carbapenem biosynthesis -

PT . useful for isolating other carbapenem synthesising genes and

PT producing carbapenem in heterologous organisms

PS Disclosure; Fig 4; 31pp; English.

CC W14445-52 are the protein products of the car genes (carA-H) which are

CC involved in the biosynthesis of carbapenem, an antibiotic which acts as

CC a beta-lactamase inhibitor. The car biosynthetic genes are regulated by

CC the carR gene product (W14444), a positive activator of the biosynthetic

CC genes which functions in trans. The car genes can be used to identify

CC homologous genes in bacteria and fungi, while products of the genes may

CC be used to raise antibodies used to detect clones (in a *Streptomyces*

CC library) that express cross-reactive protein.

SQ Sequence 521 AA;

```

Db      64 kcemergtayligslynrtfliglagvwegeay 96
        || :: | :||: | : || : ||| |
Qy      34 KC-VQAGVTYLFVQLCKMLFLATFFPTWEGGIY 65

```

```

RESULT      9
ID   W02722 standard; peptide; 349 AA.
AC   W02722;
DT   13-NOV-1996 (first entry)
DE   G-protein coupled mouse glucocorticoid-induced receptor.
KW   G-protein coupled receptor; ligand binding assay; transmembrane domain;
KW   schizophrenia; dopamine; cAMP; adenosine; thrombin; adrenergic; opsin;
KW   muscarinic acetylcholine; endothelin; bombesin; endocrine; rhodopsin;
KW   odorant; cytomegalovirus; serotonergic.
OS   Mus musculus.
PN   US5508384-A.
PD   16-APR-1996.
PF   10-SEP-1992; 943236.
PR   10-SEP-1992; US-943236.
PR   09-SEP-1993; US-118270.
PA   (UJNY ) UNIV NEW YORK STATE.
PI   Murphy RB, Schuster DI;
DR   WPI; 96-208785/21.
PT   New dopamine receptor peptide - useful as antipsychotic agent, e.g.
PT   for treating schizophrenia
PS   Disclosure; Column 163-166; 184pp; English.
CC   Proteins W02657-W02730 represent a range of G-protein coupled receptor
CC   (GPR) proteins selected from cAMP, adenosine, muscarinic acetylcholine,
CC   adrenergic, thrombin, endothelin, bombesin, endocrine, rhodopsin, opsin,
CC   odorant, cytomegaloviral and other GPR proteins. The receptor proteins
CC   were used to design polypeptides, pref. based on the transmembrane
CC   domains, for use in G-protein coupled receptor ligand binding assays.
CC   The polypeptide fragments retain biological activity such as binding a
CC   GPR ligand or modulating GPR ligand binding to a GPR (see W02747-W02999
CC   for examples of polypeptide fragments). The polypeptide fragments can
CC   be used in compositions for treating subjects suffering from a pathology
CC   related to a GPR abnormality e.g. a psychotic disorder such as
CC   schizophrenia.
SQ   Sequence 349 AA;

```

Db 94 aiavdhrqhwimlpkprisitkgviyiaivwmffslphaicqklftfkysedivrslcl 153  
:  
| | : : : : || : : | :|| : | : : : : : : : : : :  
Qy 125 GIEFDWKYIQMSIDSNISLVHYIVASAOVWMITRYDLYHNFRAVLLLMFL-SVYKAFVM 183

Db 154 dpfpepadl-fwkyldiatfillyllpl 180  
: | | | : :: ||:  
Qy 184 ETFVHCLSLGSGWARLDARA-VVTGLLAL 210

```

RESULT 10
ID R48750 standard; Protein; 349 AA.
AC R48750;
DT 07-JUN-1996 (first entry)
DE G-protein coupled mouse glucocorticoid-induced receptor protein.
KW G-protein coupled receptor; ligand binding assay; transmembrane domain;
KW psychotic disorder; schizophrenia; dopamine; cAMP; adenosine; thrombin;
KW muscarinic acetylcholine; adrenergic; endothelin; bombesin; endocrine;
KW rhodopsin; opsin; odorant; cytomegalovirus.
OS Mus musculus.
PN W09405695-A1.
PD 17-MAR-1994.
PF 09-SEP-1993; U08528.
PR 10-SEP-1992; US-943236.
PA (UINY ) UNIV NEW YORK STATE.
PI Murphy RB, Schuster DI;

```

```

RESULT 13
ID R79929 standard; Protein; 417 AA.
AC R79929;
DT 09-MAY-1996 (first entry)
DE Porcine acylglucosamine-2-epimerase mutant.
KW Porcine; acylglucosamine-2-epimerase; N-acetylmannosamine;
KW N-acetylneuraminic acid; renin-binding; enzymatic production;
KW mutant.
OS Sus scrofa.
FH Key Location/Qualifiers
FT misc_difference 23
FT /note= "wild type Met substd. with Val"
FT misc_difference 27
FT /note= "wild type Leu substd. with Met"
FT misc_difference 33
FT /note= "wild type Arg substd. with Gln"
FT misc_difference 45
FT /note= "wild type Asp substd. with Glu"
FT misc_difference 71
FT /note= "wild type Lys substd. with Thr"
FT misc_difference 72
FT /note= "wild type Leu substd. with Phe"
FT misc_difference 76
FT /note= "wild type His substd. with Arg"
FT misc_difference 77
FT /note= "wild type Arg substd. with His"
FT misc_difference 78
FT /note= "wild type Pro substd. with Ala"
FT misc_difference 79
FT /note= "wild type Glu substd. with Gln"
FT misc_difference 94
FT /note= "wild type His substd. with Tyr"
FT misc_difference 101
FT /note= "wild type Glu substd. with Gly"
FT misc_difference 120
FT /note= "wild type Ser Substd. with Thr"
FT misc_difference 137
FT /note= "wild type Val substd. with Ala"
FT misc_difference 139
FT /note= "wild type Ala substd. with Gly"
FT misc_difference 141
FT /note= "wild type Ala substd. with Val"
FT misc_difference 145

```

Db     269 wvgyr-ysi-wsaiglsialhvqfdhfegllsgahwmdqhfk-lkptleknappvllallgi 325  
             ||| | : | | : || :::: |: : | || : | : : | : || ::  
Qy     120 WVGARGIEFDWKYIQMSIDSNSISLVHYI-VASAOVVMITRYDLYHNF-RPAVLMLFSLV 177

Db 326 wy 327  
Qy 178 -Y 178

## RESULT 15

ID W48271 standard; Protein; 152 AA.  
AC W48271;  
DT 23-JUN-1998 (first entry)  
DE Rat ninjurin 1.  
KW Rat; ninjurin; cellular adhesion molecule; membrane bound; tumour;  
KW nerve injury induced gene; inflammation; nervous system.  
OS Rattus sp.  
PN W09803650-A1.  
PD 29-JAN-1998.  
PF 24-JUL-1997; U12210.  
PR 24-JUL-1996; US-672850.  
PA (UNIW ) UNIV WASHINGTON.  
PI Araki T, Milbrandt J;  
DR WPI; 98-120775/11.  
DR N-PSDB; V20665.  
PT New isolated nerve injury induced (ninjurin) gene - used to develop  
PT products for treating conditions involving excessive or insufficient  
PT cellular adhesion, e.g. inflammation or tumours  
PS Claim 5; Fig 1B; 80pp; English.  
CC The present sequence represents rat ninjurin 1 (nerve injury induced).  
CC The ninjurin protein (NP) plays a role in axonal regeneration of  
CC peripheral nervous system (PNS) neuronal cells after injury. The  
CC products can be used for developing products for treating ninjurin  
CC mediated disorders including conditions involving inappropriate (i.e.  
CC excessive or insufficient) cellular adhesion. Conditions involving  
CC excessive cellular adhesion which may be treated include e.g.  
CC inflammatory diseases such as rheumatoid arthritis, asthma, allergy  
CC conditions, adult respiratory distress syndrome, inflammatory bowel  
CC diseases (e.g. Crohn's disease, ulcerative colitis and regional  
CC enteritis) and ophthalmic inflammatory diseases, autoimmune diseases,  
CC thrombosis or inappropriate platelet aggregation conditions,  
CC arteriosclerosis, reocclusion following thrombolysis, cardiovascular  
CC diseases, some forms of diabetes and neoplastic disease including  
CC metastasis conditions. The products can also be used to promote  
CC cellular adhesion, e.g. in nerve regeneration, wound healing or  
CC prosthetic implantation. The products can also be used for detection,  
CC purification, diagnosis and screening assays.  
SQ Sequence 152 AA;

Query Match 4.4%; Score 85; DB 29; Length 152;  
Best Local Similarity 27.6%; Pred. No. 1.05e+02;  
Matches 21; Conservative 19; Mismatches 31; Indels 5; Gaps 5;

Db 56 llmanasqlkavveqgnef-fvplvllisisvlqigvgvlliflv-kydlnnpakha 113  
|:|: | | :| || : : : :| | | | : : : :| | : : |  
Qy 111 LIMSRCIPLW-VGARGIEFDWKYIQMSIDSNISLVHYI-VASAQVVMITRYDLYHFRPA 168  
Db 114 kldflnnlatglvfii 129  
| : | : | : :  
Qy 169 VL-LLMFLSVYKAFVM 183

Search completed: Fri Jan 22 17:59:45 1999  
Job time : 84 secs.

\*\*\*\*\*

MSRNA (TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Fri Jan 22 18:01:16 1999; MasPar time 15.58 Seconds  
782.839 Million cell updates/sec

Tabular output not generated.

Title: >US-09-162-597-3  
Description: (1-245) from US09162597.pep  
Perfect Score: 1940  
Sequence: 1 MTLFHFNCALAFYFPYFIT.....GLVSTLMLYFPASLQVLVK 245

Scoring table: PAM 150  
Gap 11

Searched: 165420 seqs, 49795644 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: sptrembl6  
1:sp\_archaea 2:sp\_bacteria 3:sp\_fungi 4:sp\_human  
5:sp\_invertebrate 6:sp\_mammal 7:sp\_mhc 8:sp\_organelle  
9:sp\_phage 10:sp\_plant 11:sp\_rodent 12:sp\_unclassified  
13:sp\_vertebrate 14:sp\_virus

Statistics: Mean 46.110; Variance 101.374; scale 0.455

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	No.	Score	Match	Length	DB	ID	Description	Pred. No.
1	636	32.8	216	5	Q23483	COSMID ZK418.	2.66e-90	
2	119	6.1	1087	5	Q46089	1-EVIDENCE-PREDICTED B	1.50e-02	
3	113	5.8	215	8	Q36712	PETB, PETD {RNA EDITIN	8.40e-02	
4	113	5.8	232	8	Q32438	POT. ALT. PETB GENE PR	8.40e-02	
5	113	5.8	232	8	Q33298	ALTERNATE PETB GENE PR	8.40e-02	
6	113	5.8	232	8	Q36615	APOCYTOCHROME B6.	8.40e-02	
7	109	5.6	300	2	Q32730	ATTA2 PROTEIN.	2.58e-01	
8	106	5.5	271	8	Q63292	NADH DEHYDROGENASE SUB	5.88e-01	
9	105	5.4	212	1	Q27297	CONSERVED PROTEIN.	7.72e-01	
10	104	5.4	310	2	Q51297	CONSERVED HYPOTHETICAL	1.01e+00	
11	104	5.4	379	8	Q47977	CYTOCHROME B.	1.01e+00	
12	105	5.4	527	10	Q64515	YUP8H12R.2 PROTEIN.	7.72e-01	
13	104	5.4	1410	5	Q18259	SIMILARITY ALONG ENTIR	1.01e+00	
14	103	5.3	308	5	Q17615	C29F3.6.	1.32e+00	
15	102	5.3	379	8	Q47963	CYTOCHROME B.	1.73e+00	
16	102	5.3	379	8	Q47962	CYTOCHROME B.	1.73e+00	
17	102	5.3	380	8	Q48000	CYTOCHROME B (FRAGMENT	1.73e+00	
18	102	5.3	503	8	Q37612	CYTOCHROME C OXIDASE'S	1.73e+00	
19	101	5.2	204	8	Q47043	CYTOCHROME B (FRAGMENT	2.25e+00	
20	101	5.2	348	8	Q47403	NADH DEHYDROGENASE SUB	2.25e+00	

21	100	5.2	348	8	Q47399	NADH DEHYDROGENASE SUB	2.92e+00
22	100	5.2	348	8	Q47398	NADH DEHYDROGENASE SUB	2.92e+00
23	101	5.2	379	8	Q47991	CYTOCHROME B.	2.25e+00
24	101	5.2	379	8	Q47965	CYTOCHROME B.	2.25e+00
25	101	5.2	379	8	Q47992	CYTOCHROME B.	2.25e+00
26	101	5.2	379	8	Q47972	CYTOCHROME B.	2.25e+00
27	101	5.2	379	8	Q47966	CYTOCHROME B.	2.25e+00
28	101	5.2	379	8	Q47974	CYTOCHROME B.	2.25e+00
29	101	5.2	379	8	Q47970	CYTOCHROME B.	2.25e+00
30	100	5.2	379	8	Q47982	CYTOCHROME B.	2.92e+00
31	101	5.2	380	8	Q47971	CYTOCHROME B.	2.25e+00
32	101	5.2	380	8	Q47976	CYTOCHROME B.	2.25e+00
33	101	5.2	380	8	Q47964	CYTOCHROME B.	2.25e+00
34	101	5.2	380	8	Q47990	CYTOCHROME B.	2.25e+00
35	101	5.2	380	8	Q47975	CYTOCHROME B.	2.25e+00
36	101	5.2	495	8	Q34463	CYTOCHROME OXIDASE SUB	2.25e+00
37	101	5.2	550	2	Q31692	PHOSPHOTRANSFERASE SYS	2.25e+00
38	101	5.2	573	2	Q69251	ENZYME I.	2.25e+00
39	99	5.1	133	8	Q47442	CYTOCHROME B (FRAGMENT	3.80e+00
40	99	5.1	247	1	Q50081	247AA LONG HYPOTHETICA	3.80e+00
41	99	5.1	345	5	Q20963	RELATED TO C. ELEGANS	3.80e+00
42	99	5.1	363	1	Q26434	PEROSAMINE SYNTHETASE.	3.80e+00
43	99	5.1	379	8	Q48001	CYTOCHROME B.	3.80e+00
44	99	5.1	379	8	Q48364	CYTOCHROME B.	3.80e+00
45	99	5.1	2206	5	P91744	VOLTAGE-DEPENDENT CALC	3.80e+00

#### ALIGNMENTS

RESULT 1

ID Q23483 PRELIMINARY; PRT; 216 AA.

AC Q23483;

DT 01-NOV-1996 (TREMBLREL. 01, CREATED)

DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)

DT 01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)

DE COSMID ZK418.

GN ZK418.5.

OS CAENORHABDITIS ELEGANS.

OC EUKARYOTA; METAZOA; ACOELOMATES; NEMATODA; SECERNENTEA; RHABDITIDA.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2;

RX MEDLINE; 94150718.

RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M., BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A., CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L., GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L., JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P., LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M., PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R., SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J., THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R., WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;

RL NATURE 368:32-38(1994).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2;

RA FULTON L.;

RL SUBMITTED (APR-1994) TO EMBL/GENBANK/DBJ DATA BANKS.

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2;

RA WATERSTON R.;

RL SUBMITTED (APR-1994) TO EMBL/GENBANK/DBJ DATA BANKS.

DR EMBL; U00047; G470373; -.

SQ SEQUENCE 216 AA; 23931 MW; 8B1C508A CRC32;

Query Match 32.8%; Score 636; DB 5; Length 216;  
Best Local Similarity 47.8%; Pred. No. 2.66e-90;  
Matches 76; Conservative 41; Mismatches 41; Indels 1; Gaps 1;

Db 1 MSFHFHNCALAFYFIVYKYSINEYSSIWKCATASGGYLLTQLAKLLIATFFPAL 60  
|::||| |||||: |||| || ||::||::||| |::||| ||::|||::|



.RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DELTA H;  
 RX MEDLINE; 98037514.  
 RA SMITH D.R., DOUCETTE-STAMM L.A., DELOUGHERY C., LEE H.-M., DUBOIS J.,  
 RA ALDREDGE T., BASHIRZADEH R., BLAKELY D., COOK R., GILBERT K.,



Query Match 5.4%; Score 105; DB 10; Length 527;

Qy : 182 v 182

Search completed: Fri Jan 22 18:02:50 1999  
Job time : 94 secs.

US-09-162-597-3

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☒ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☒ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☒ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☒ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:**

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**